

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

EDITED BY

CARL MURCHISON

Gordon Allport  
Harvard University

Arthur G. Bills  
University of Cincinnati

E. A. Bott  
University of Toronto

Leonard Carmichael  
Smithsonian Institution

Herbert S. Conrad  
U. S. Office of Education

Elmer K. Culler  
University of Rochester

John F. Dashiell  
University of North  
Carolina

Roy M. Dorcus  
University of California  
at Los Angeles

Horace B. English  
Ohio State University

Franklin Fearing  
University of California  
at Los Angeles

Frank A. Geldard  
University of Virginia

Clarence H. Graham  
Columbia University

Ernest R. Hilgard  
Stanford University

H. M. Johnson  
Tulane University

Heinrich Klüver  
University of Chicago

Carney Landis  
Columbia University

K. S. Lashley  
Harvard University

Norman L. Munn  
Bowdoin College

John Paul Nafe  
Washington University

Donald G. Paterson  
University of Minnesota

Lee Edward Travis  
University of Southern  
California

Robert C. Tryon  
University of California

Michael J. Zigler  
Wellesley College

If this space should be unstamped, this is the regular library edition.  
But if this space is stamped with a designating title, this is a special

edition, sold under the restrictions of a bilateral contract, and may not  
be resold for a period of five years from date of publication.

## VOLUME 39

1955

Copyright, 1955, by The Journal Press  
Published Quarterly by The Journal Press  
Provincetown, Massachusetts, U. S. A.



# AUTHOR INDEX

Abramson, H. A. ....	3, 127, 373, 435, 443, 455, 465	Kornetsky, C. ....	3
Alexander, S. J. ....	389, 403, 411	Lachman, Sheldon J. ....	325
Angelino, Henry ....	195, 321	Landis, Carney ....	183
Atkinson, W. H. ....	371	Lepley, William M. ....	215
Ballin, H. M. ....	77	Levine, A. ....	3
Barratt, Ernest S. ....	279	Lofchie, H. ....	109
Beebe-Center, J. G. ....	157, 371	Mech, Edmund V. ....	195, 321
Bernberg, Raymond E. ....	89	Meer, Bernard ....	117
Cobliner, W. Godfrey ....	253	Miner, John B. ....	199
Cotzin, M. ....	389, 403, 411, 417	Muthard, John E. ....	269
Crannell, C. W. ....	351	Noble, Clyde E. ....	475
Darby, C. L. ....	105	O'Connell, D. N. ....	157
Diamond, Solomon ....	385	Plaut, Edika ....	351
Ewald, A. T. ....	435, 465	Pressey, S. L. ....	485
Farese, Francis J. ....	475	Prothro, E. Terry ....	247
Gellhorn, E. ....	77	Ricciuti, E. A. ....	403
Ghent, Lila ....	227	Riopelle, A. J. ....	105
Green, Russel F. ....	355	Rogers, M. S. ....	157, 371
Hamwi, Violet ....	183	Rollin, A. Robert ....	301
Heidbreder, Edna ....	341	Sapon, Stanley M. ....	97
Henkin, Robert I. ....	161	Secord, Paul F. ....	269
Hill, C. J., Jr. ....	403, 417	Semmes, Josephine ....	227
Hirsch, M. W. ....	373, 435, 443, 455, 465	Stein, Morris I. ....	117
Jarvik, M. E. ....	3, 373, 435, 443, 455, 465	Stern, Ferdinand ....	71
Johnson, Carson ....	423	Stotsky, Bernard A. ....	325
Jones, A. W. ....	485	Teuber, Hans-Lukas ....	227
Juzak, Tatania ....	329	Tomkins, Silvan S. ....	199
Kaden, Stanley E. ....	61	Voeks, Virginia W. ....	289
Karson, Samuel ....	71	Wagner, M. ....	3
Kaufman, M. R. ....	3	Wapner, Seymour ....	61
Klee, J. B. ....	411	Weinstein, Sidney ....	227
Koella, W. P. ....	77	Wendt, G. R. ....	389, 403, 411, 417, 423
		Werner, Heinz ....	61

Encl. Rev. Research

11. 8. 70

9113



## TABLE OF CONTENTS

Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses BY H. A. ABRAMSON, M. E. JÄRVIK, M. R. KAUFMAN, A. LEVINE, C. KORNETSKY, AND M. WAGNER	3
Studies in physiognomic perception: II. Effect of directional dynamics of pictured objects and of words on the position of the apparent horizon BY STANLEY E. KADEN, SEYMOUR WAPNER, AND HEINZ WERNER	61
A critique of Mowrer's theory of neurosis* BY FERDINAND STERN AND SAMUEL KARSON	71
The influence of hypothalamic stimulation on evoked cortical potentials BY E. GELLHORN, W. P. KOELLA, AND H. M. BALLIN	77
A measure of social conformity BY RAYMOND E. BERNBERG	89
A work sample test for foreign language prognosis BY STANLEY M. SAPON	97
Differential problem sequences and the formation of learning sets BY C. L. DARBY AND A. J. RIOPELLE	105
The performance of adults under distraction stress: A developmental approach BY STANLEY H. LOFCHIE	109
Measures of intelligence and creativity BY BERNARD MEER AND MORRIS I. STEIN	117
Lysergic acid diethylamide (LSD-25): III. As an adjunct to psychotherapy with elimination of fear of homosexuality BY HAROLD A. ABRAMSON	127
Transmission of information about sucrose and saline solutions through the sense of taste BY J. G. BEEBE-CENTER, M. S. ROGERS, AND D. N. O'CONNELL	157
A factorial study of the components of music BY ROBERT I. HENKIN	161
Memory for color BY VIOLET HAMWI AND CARNEY LANDIS	183
"Fears and worries" concerning physical changes: A preliminary survey of 32 females BY HENRY ANGELINO AND EDMUND V. MECH	195
Contributions to the standardization of the Tomkins-Horn Picture Arrangement Test: Plate norms BY SILVAN S. TOMKINS AND JOHN B. MINER	199
The rationale, construction, and preliminary try-out of the synonym vocabulary test BY WILLIAM M. LEPLEY	215
Spatial orientation in man after cerebral injury: Analyses by locus of lesion BY JOSEPHINE SEMMES, SIDNEY WEINSTEIN, LILA GHENT, AND HANS-LUKAS TEUBER	227
An alternative approach in cross-cultural intelligence testing BY E. TERRY PROTHRO	247
Intra-communication and attitude: A methodological note BY W. GODFREY COBLINER	253
Personalities in faces: IV. A descriptive analysis of the perception of women's faces and the identification of some physiognomic determinants BY PAUL F. SECORD AND JOHN E. MUTHARD	269
The space-visualization factors related to temperament traits BY ERNEST S. BARRATT	279

\*On p. 72, 1.27, the word *legal* should be replaced by the words *popular* (psychological).

Gradual strengthening of S-R connections or increasing number of S-R connections . . . . .	289
By VIRGINIA W. VOEKS . . . . .	
Mass action effects in learning . . . . .	301
By A. ROBERT ROLLIN . . . . .	
Some "first" sources of sex information as reported by sixty-seven college women . . . . .	321
By HENRY ANGELINO AND EDMUND V. MECH . . . . .	
The authoritarian personality as a stereotype . . . . .	325
By BERNARD A. STOTSKY AND SHELDON J. LACHMAN . . . . .	
The effects of praise and reproof on the generalization of learned concepts . . . . .	329
By TATANIA JUZAK . . . . .	
Stimulus-discriminability and concept-attainment: A question arising from Baum's experiment . . . . .	341
By EDNA HEIDBREDER . . . . .	
Drawings of a three-dimensional object by mental patients: A preliminary report . . . . .	351
By C. W. CRANNELL AND EDIKA PLAUT . . . . .	
Transfer of skill on a following tracking task as a function of task difficulty (target size) . . . . .	355
By RUSSEL F. GREEN . . . . .	
Intensive equivalences for sucrose and NaCl solutions . . . . .	371
By J. G. BEEBE-CENTER, M. S. ROGERS, AND W. H. ATKINSON . . . . .	
Lysergic acid diethylamide (LSD-25): IV. Effect on attention and concentration . . . . .	373
By M. E. JARVIK, H. A. ABRAMSON, AND M. W. HIRSCH . . . . .	
Sex stereotypes and acceptance of sex rôle . . . . .	385
By SOLOMON DIAMOND . . . . .	
Chemical influences on behavior: I. The effects of a small dose of hyoscine on performance . . . . .	389
By S. J. ALEXANDER, M. COTZIN, AND G. R. WENDT . . . . .	
Studies of motion sickness: X. Experimental proof that aviation cadets tell the truth on motion sickness history questionnaires . . . . .	403
By S. J. ALEXANDER, M. COTZIN, C. J. HILL, JR., E. A. RICCIUTI, AND G. R. WENDT . . . . .	
Studies of motion sickness: XIII. The effects of sickness upon rifle target shooting . . . . .	411
By S. J. ALEXANDER, M. COTZIN, J. B. KLEE, AND G. R. WENDT . . . . .	
Studies of motion sickness: XIV. Subjective reports of the apparent path of motion on a vertical accelerator . . . . .	417
By M. COTZIN, C. J. HILL, JR., AND G. R. WENDT . . . . .	
Studies of motion sickness: XVII. The effects of temperature, posture, and wave frequency upon sickness rates . . . . .	423
By CARSON JOHNSON AND G. R. WENDT . . . . .	
Lysergic acid diethylamide (LSD-25): V. Effect on spatial relations abilities . . . . .	435
By H. A. ABRAMSON, M. E. JARVIK, M. W. HIRSCH, AND A. T. EWALD . . . . .	
Lysergic acid diethylamide (LSD-25): VI. Effect upon recall and recognition of various stimuli . . . . .	443
By M. E. JARVIK, H. A. ABRAMSON, AND M. W. HIRSCH . . . . .	
Lysergic acid diethylamide (LSD-25): VII. Effect upon two measures of motor performance . . . . .	455
By H. A. ABRAMSON, M. E. JARVIK, AND M. W. HIRSCH . . . . .	
Lysergic acid diethylamide (LSD-25): VIII. Effect on arithmetic test performance . . . . .	465
By M. E. JARVIK, H. A. ABRAMSON, M. W. HIRSCH, AND A. T. EWALD . . . . .	
An apparatus for research in human selective learning . . . . .	475
By CLYDE E. NOBLE AND FRANCIS J. FARESE . . . . .	
1923-1953 and 20-60 age changes in moral codes, anxieties, and interests, as shown by the "X-O Tests" . . . . .	485
By S. L. PRESSEY AND A. W. JONES . . . . .	



\$20.00 per annum  
\$10.00 per volume  
\$5.00 per half volume

QUARTERLY  
Two volumes per year  
Immediate publication

January, 1955  
Volume 39, First Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

JANUARY, 1955

(Manuscripts are printed in the order of final acceptance)

Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses . . . . .	3
By H. A. ABRAMSON, M. E. JARVIK, M. R. KAUFMAN, A. LEVINE, C. KORNETSKY, AND M. WAGNER	
Studies in physiognomic perception: II. Effect of directional dynamics of pictured objects and of words on the position of the apparent horizon . . . . .	61
By STANLEY E. KADEN, SEYMOUR WAPNER, AND HEINZ WERNER	
A critique of Mowrer's theory of neurosis . . . . .	71
By FERDINAND STERN AND SAMUEL KARSON	
The influence of hypothalamic stimulation on evoked cortical potentials . . . . .	77
By E. GELLHORN, W. P. KOELLA, AND H. M. BALLIN	
A measure of social conformity . . . . .	89
By RAYMOND E. BERNBERG	
A work sample test for foreign language prognosis . . . . .	97
By STANLEY M. SAPON	

(OVER)

Copyright, 1955, by The Journal Press  
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

Differential problem sequences and the formation of learning sets . . . .	105
By C. L. DARBY AND A. J. RIOPELLE	
The performance of adults under distraction stress: A developmental approach	109
By STANLEY H. LOFCHIE	
Measures of intelligence and creativity . . . . .	117
By BERNARD MEER AND MORRIS I. STEIN	
Lysergic acid diethylamide (LSD-25): III. As an adjunct to psychotherapy with elimination of fear of homosexuality . . . . .	127
By HAROLD A. ABRAMSON	
Transmission of information about sucrose and saline solutions through the sense of taste . . . . .	157
By J. G. BEEBE-CENTER, M. S. ROGERS, AND D. N. O'CONNELL	
A factorial study of the components of music . . . . .	161
By ROBERT I. HENKIN	
Memory for color . . . . .	183
By VIOLET HAMWI AND CARNEY LANDIS	
"Fears and worries" concerning physical changes: A preliminary survey of 32 females . . . . .	195
By HENRY ANGELINO AND EDMUND V. MECH	
Contributions to the standardization of the Tomkins-Horn Picture Arrange- ment Test: Plate norms . . . . .	199
By SILVAN S. TOMKINS AND JOHN B. MINER	
<sup>n</sup> The rationale, construction, and preliminary try-out of the synonym vocabu- lary test . . . . .	215
By WILLIAM M. LEPLEY	
Spatial orientation in man after cerebral injury: Analyses by locus of lesion	227
By JOSEPHINE SEMMES, SIDNEY WEINSTEIN, LILA GHENT, AND HANS- LUKAS TEUBER	



## LYSERGIC ACID DIETHYLAMIDE (LSD-25): I. PHYSIOLOGICAL AND PERCEPTUAL RESPONSES\*

*Departments of Medicine, Neurology, and Psychiatry, Mount Sinai Hospital, New York City*

H. A. ABRAMSON, M. E. JARVIK, M. R. KAUFMAN, C. KORNETSKY,  
A. LEVINE, AND M. WAGNER

### A. INTRODUCTION

The purpose of this report is to present the variety and course of physiological and perceptual symptoms subjectively reported by "normal" human subjects reacting to various doses of lysergic acid diethylamide. The responses were elicited by means of a questionnaire containing 47 items. These items were compiled by going through the literature and listing the reported symptoms and signs, ostensibly due to LSD-25. The actual wording of the items in the questionnaire is listed in Table 2.

Early in our experiments it was discovered that certain subjects may give positive responses when only a placebo (tap water) was administered. In these subjects the scores on the questionnaire would make it appear that these subjects had actually been given LSD-25 instead of nothing at all. Indeed, the power of suggestion is so great that one of us (H.A.A.) has noted that an observer of the experiments reported more positive responses than the subject who had actually received LSD-25. To illustrate the nature of the responses under zero dosage of LSD-25, Table 1 summarizes five typical responses of subjects under our experimental conditions. Table 1A illustrates part of a questionnaire to aid the reader in understanding the way the items were scored. This was from Subject 12-64A (zero dosage). In the table, numbers 1-47 represent the questions of the questionnaire listed in Table 2. The response to the questionnaire as a function of time is shown for five separate subjects: 31-73B, 22-62C, 35-86B, 20-58B, and 12-64A. The experiments were performed as part of a group test. Note in the table that Subject 31-73B acted as an ideal subject. There were no positive responses to zero dosage during the entire experimental period of more than five hours. Subject 22-62C also shows only one positive response at one-half hour. In the case of Subject 35-86B the picture changes. There are now 12 different positive re-

\*Received in the Editorial Office on July 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

TABLE 1  
POSITIVE RESPONSES TO QUESTIONNAIRE BY SELECTED SUBJECTS  
UNDER ZERO DOSE OF LSD-25

(Five columns represent five separate subjects. Time, in hours after placebo, is indicated by numbers  $\frac{1}{2}$  -  $3\frac{1}{2}$ .)

Question number	31-73B					22-62C					35-86B					20-58B					12-64A				
	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$
1																									x
2																									x
3											x											x			
4																						x	x	x	x
5																									
6																	x					x	x		
7																									
8																									
9																	x	x				x			x
10																									x
11																									
12																									
13											x	x	x									x	x	x	x
14																									
15											x														
16											x						x								
17																									
18																									
19																									
20																	x								
21																									
22																									



TABLE 1 (continued)

Question number	31-73B					22-62C					35-86B					20-58B					12-64A				
	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{2}$	$3\frac{1}{2}$	$3\frac{1}{2}+$
23											x	x	x												
24						x										x		x				x			
25																									
26																									
27																									
28																						x		x	
29																						x		x	
30																						x		x	
31																						x		x	
32											x	x	x									x		x	
33																						x		x	
34																						x		x	
35											x	x	x									x		x	
36																									
37																									
38																									
39																									
40																									
41																									
42																						x		x	
43																									
44																									
45											x														
46																						x		x	
47											x	x	x	x	x							x	x	x	x

TABLE 1a

TYPICAL RESPONSE TO QUESTIONNAIRE: BLANK SPACES INDICATE NEGATIVE RESPONSES (Half an hour after administration and at hourly intervals thereafter, the following direct questions are put to the subject. His rating is accepted, if possible on a + to +++++ basis)

Question	Up to ½ Hr.	Up to 1½ Hrs.	Up to 2½ Hrs.	Up to 3½ Hrs.
1. Do you feel ill in any way?	.	.	.	+
2. Are you nauseated?	.	.	.	+
3. Have you a feeling of choking?	+	.	.	.
4. Is salivation increased?	++	++	+	++
5. Or decreased?	.	+	.	.
6. Is your appetite increased?	+	.	+	.
7. Or decreased?	.	++	.	.
8. Do you have a "dry" taste in your mouth?	.	.	.	.
9. Do you have a funny taste in your mouth?	+	.	.	++
10. Is it a bitter taste?	.	.	.	++
11. Are your lips numb?	.	.	.	.
12. Or drawn back as if you were smiling?	.	.	.	.
13. Does your head ache?	+	+	+	++++
14. Are things moving around you?	.	.	.	.
15. Do you feel dizzy?	.	.	.	.
16. Or unsteady?	.	.	.	.
17. Is there difficulty in breathing?	.	.	.	.
18. Do you pass more urine than usual?	.	.	.	.
19. Are you aware of your heart beat?	.	.	.	.
20. Is it faster than usual?	.	.	.	.
21. Are you sweating?	.	.	.	.
22. Are you hot?	.	.	.	.
23. Or cold?	.	.	.	.
24. Are your palms moist?	++	.	.	.
25. Or dry?	.	.	.	.
26. Or cold?	.	.	.	.
27. Is your skin sensitive?	+	.	+	+
28. Do you have funny feelings on your skin?	++	.	.	+
29. Do your hands and feet feel peculiar?	++	+	.	++
30. Do they feel heavy?	++	++	.	.
31. Or light?	+	++	.	++++
32. Is there pressure in your ears?	+	+	.	++
33. Is your hearing abnormal?	++	+	.	++
34. Is it more acute than usual?	.	.	.	++
35. Is your eyesight blurred?	.	.	.	.
36. Do you have difficulty in focusing your vision?	.	.	.	.
37. Do you see double?	.	.	.	.
38. Are shapes and colors altered in any way?	.	++	.	+
39. Does light bother you?	.	.	.	.
40. Do things seem too close?	.	.	.	.
41. Or too far away?	.	.	.	.
42. Do you tremble inside?	++	+	.	++++
43. Do you feel weak?	.	.	.	.
44. Or fatigued?	.	.	.	.
45. Do you feel drowsy?	.	.	.	.
46. Do you feel as if in a dream?	.	.	.	.
47. Are you anxious?	+	+	+	++
	15	12	5	15



TABLE 2  
AVERAGE PERCENTS OF SUBJECTS IN THREE LSD-25 DOSAGE GROUPS RESPONDING  
POSITIVELY TO EACH QUESTION

Question Number	Question	Percent of subjects		
		N = 13 0 Gamma	N = 21 25-75 Gamma	N = 14 100-225 Gamma
1	Do you feel ill in any way?	8	11	24
2	Are you nauseated?	3	15	26
3	Have you a feeling of choking?	1	11	12
4	Is salivation increased?	9	15	6
5	Or decreased?	6	12	22
6	Is your appetite increased?	5	11	17
7	Or decreased?	11	14	19
8	Do you have a "dry" taste in your mouth?	5	17	16
9	Do you have a funny taste in your mouth?	8	19	19
10	Is it a bitter taste?	4	7	9
11	Are your lips numb?	1	16	21
12	Or drawn back as if you were smiling?	0	19	14
13	Does your head ache?	20	27	20
14	Are things moving around you?	0	9	17
15	Do you feel dizzy?	7	29	40
16	Or unsteady?	3	43	48
17	Is there difficulty in breathing?	3	9	7
18	Do you pass more urine than usual?	5	7	6
19	Are you aware of your heart beat?	2	17	7
20	Is it faster than usual?	5	10	8
21	Are you sweating?	1	17	18
22	Are you hot?	7	26	21
23	Or cold?	0	9	8
24	Are your palms moist?	35	45	41
25	Are your palms dry?	0	3	7
26	Or cold?	9	7	7
27	Is your skin sensitive?	4	11	18
28	Do you have funny feelings on your skin?	2	31	38
29	Do your hands and feet feel peculiar?	9	41	40
30	Do they feel heavy?	7	25	30
31	Or light?	2	23	24
32	Is there pressure in your ears?	4	20	24
33	Is your hearing abnormal?	4	12	3
34	Is it more acute than usual?	4	7	15
35	Is your eyesight blurred?	2	22	24
36	Do you have difficulty in focusing your vision?	0	14	27
37	Do you see double?	0	0	2
38	Are shapes and colors altered in any way?	1	11	14

TABLE 2 (continued)

Question Number	Question	Percent of subjects		
		N = 13 0 Gamma	N = 21 25-75 Gamma	N = 14 100-225 Gamma
39	Does light bother you?	2	8	6
40	Do things seem too close?	0	7	7
41	Or too far away?	0	8	13
42	Do you tremble inside?	6	40	47
43	Do you feel weak?	2	39	41
44	Or fatigued?	11	39	23
45	Do you feel drowsy?	24	34	28
46	Do you feel as if in a dream?	6	37	33
47	Are you anxious?	13	22	27

sponses, some of which were maintained during the entire period of the experiment even though no medication at all had been administered. Note especially that Subject 35-86B had responses especially characteristic of LSD-25 itself, e.g., eyesight blurring, seeing double, and visual distortions. The pattern becomes even more marked with Subject 20-58B, who throughout the experiment felt as if he were in a dream and had peculiar sensations in his hands and feet. Perhaps most striking of all of the group presented here is Subject 12-64A who responded positively to almost half of the items and whose symptoms persisted for approximately 10 hours. Indeed, Subject 12-64A had such a severe response to the placebo that considerable care was required to maintain an experimental situation that was not traumatic.

In view of the data presented in Table 1 it appears evident that previous studies reporting the effects of LSD-25 upon "normal" subjects need re-evaluation in terms of placebo (zero) dosage. Reference may be made to Rinkel, DeShon, and Solomon's (9) work, the investigations of Stoll (12), those of Savage (11), as well as others (1, 2, 3, 4, 5, 6, 7, 8, and 10), in which the use of controls given zero dosage was not perhaps as adequate as desirable. Indeed, it was found by us that for a given group of individuals suitable evaluation of responses to LSD-25 could not be made without the use of a zero dose control group. Further, at the time that this report was compiled there were no investigations in the literature which justify the conclusion that the symptoms reported, especially at low dosage, are significantly related to LSD-25 intoxication.

## B. METHOD

### 1. General

The questionnaire, as mentioned in the Introduction, was constructed on the basis of the physiological and perceptual symptoms reported in the litera-



ture. After several revisions, the final form consisted of 47 questions so designed that a positive response indicated the presence of a given symptom (Table 2). The questions were asked by the observer one-half hour after the administration of the drug and usually at five hourly intervals thereafter.

While suggestibility could not be prevented under our conditions, it was partially controlled by having the same questions asked of each individual at each interval. Periodic questioning also enabled the evaluation of drug effects as related to time. Onset, peak, and decline of symptoms could also be fairly well determined.

Only positive and negative responses were analyzed, although graded responses were obtained.

The form of the questionnaire is illustrated in Table 1a.

## 2. *Sample*

The sample consisted of a group of 26 paid, adult volunteers. On the basis of the Rorschach, Wechsler-Bellevue Intelligence Scale, Bender Gestalt, House-Tree-Person, and Thematic Apperception tests, subjects were classified as non-psychotic or psychotic. Only those falling within the non-psychotic range were accepted as subjects. Fifteen males and 11 females were used. Ages ranged from 24 to 41 years and *IQ*'s from 100 to 137. Weights ranged from 103 to 203 pounds and height from 5 feet to 6 feet, 2 inches.

## 3. *Dose and Test Conditions*

Doses ranged from zero, in which subjects received only water, to 225 micrograms of lysergic acid diethylamide. At no time did the subjects know their dosage. Doses were divided into three categories: zero, low (25-75 micrograms LSD-25), and high (100-225 micrograms LSD-25). Thirteen subjects were tested at zero, 21 at the 25-75 microgram level, and 14 at the 100-225 microgram level. Six of the subjects received the drug at all three levels; five were tested at both the zero and 25-75 microgram level, but not at the 100-225 microgram level. Six subjects were tested at the 25-75 and 100-225 microgram levels but not at the zero level. Eight subjects were tested at only one level.

Subjects received the drug at about 10:30 or 11:00 A.M., several hours after breakfast. They were usually observed in groups of two, three, or four with the doses usually different amongst the subjects. Most of the time was spent in the experimental room where psychological and psychophysical tests were administered. Group discussions were conducted when the questionnaire was not being used.

#### 4. Analysis

Responses to the questionnaire were analyzed in terms of responses to each of the 47 items, and the number of items to which positive responses were given. Within these two categories a number of analyses were made:

##### *a. Item analysis.*

(1). The percentage of subjects in each of the three dose groups responding positively to each question at each of the six hourly intervals was determined.

(2). The average percentage of subjects in each of the three dose groups responding positively per hour, over the five-hour period, was determined.

(3). Questions were ranked within each dosage group according to average hourly percentage response. A comparison of the three rank positions of questions was made.

(4). The questions which significantly discriminate between any two of the three dose levels were determined.

##### *b. Number analysis.*

(1). The average number of positive responses given during the experimental period to the 47 questions was determined for each of three dose groups. This was done for each of six hourly intervals, separately, as well as for the total five-hour period.

(2). The differences between dosage groups in number of responses made were analyzed to determine their statistical significance.

(3). Rank positions of subjects, according to average number of positive responses made per hour at each dose level, were compared as a means of evaluating the significance of the individual subject's rank.

(4). The number of responses made by subjects tested twice at the same dose level was compared to determine the consistency of response.

(5). The subject's body weight was correlated with the number of positive responses made at a given dosage.

#### C. STATISTICAL PROCEDURE, RESULTS, AND DISCUSSION

##### 1. Item Analysis

##### *a. Hourly response to each question.*

(1). *Statistical procedure.* The 47 items on the questionnaire were subjected to analysis first by transferring the data from the questionnaire cards to IBM cards. The data were divided into three dosage groups: (a) zero micrograms LSD-25 in which there were 13 subjects who received only

water; (b) 25-75 micrograms LSD-25 in which there were 21 subjects, most of whom received 50 micrograms; and (c) 100-225 micrograms LSD-25, consisting of 14 subjects, most of whom received 100 micrograms. No subject appeared twice within any one of the three groups, although a number of the subjects appeared in two or three of the groups. In cases where a subject was tested at more than one dose within a dose group, data from the highest dose were always used. Where subjects were tested more than once at a given dose, data from the first test were used. In this way "weighting" of results was avoided.

The total number of positive responses made to each question by the subjects in each group at six hourly intervals was tabulated by means of the IBM 101 Electronic Statistical Machine. Although positive responses were originally made on a plus one to a plus five basis, it was felt that the evaluation was too subjective and the number of subjects too small to warrant separate analysis of each degree of plus response. Consequently, all plus responses were considered as positive responses and treated as equal.

The percentage responding positively each hour was then calculated by means of the following formula:

$$P = \frac{\sum_{i=1}^N X}{N} \times 100 \quad (1)$$

where  $P$  equals the percentage,  $\sum_{i=1}^N X$  equals the sum of plus responses, and  $N$

equals the number of subjects in the groups

(2). *Results.* Figure 1, containing Items 1-47, and our discussion of items 1-47, indicates the percentage of subjects in the three dose groups giving positive responses to the items at each of the hourly intervals. The percentage of subjects responding positively to each question is placed on the ordinate of a co-ordinate system, and the time of response is placed on the abscissa. Each point represents the percentage response at a given time. The time curve of each question is shown on a separate graph.

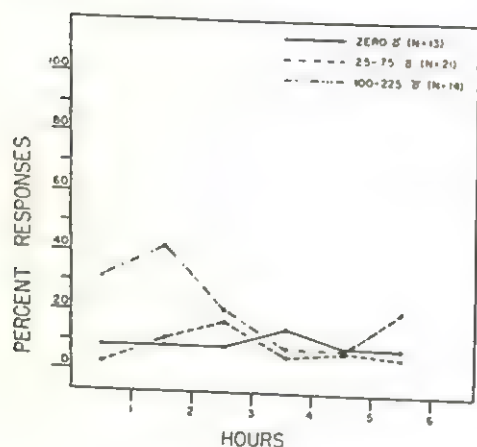
In the following paragraphs the results shown in Figure 1 are briefly discussed.

*Item 1: "Do you feel ill in any way?"* For zero dosage, out of 13 subjects only 8 per cent responded positively in each interval, except for the fourth interval where two responded positively. In terms of percentages, the peak occurred at  $3\frac{1}{2}$  hours for zero dosage with 15 per cent responding posi-

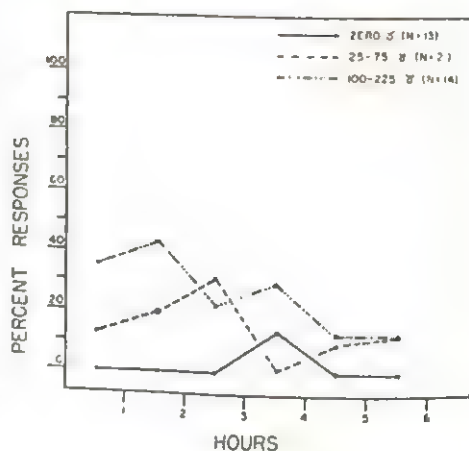


tively. For a dosage of 25-75 micrograms the peak occurred at  $2\frac{1}{2}$  hours with 19 per cent responding positively. For the dosages of 100-225 micrograms, when the peak occurred approximately  $1\frac{1}{2}$  hours after the drug, 43 per cent responded positively. Thus, there was both a more rapid and marked effect from the highest dose of LSD-25, where almost half of the people reported feeling ill in the second hour. It is interesting to see that for this high dose the percentage of people feeling ill declined very rapidly after  $1\frac{1}{2}$  hours.

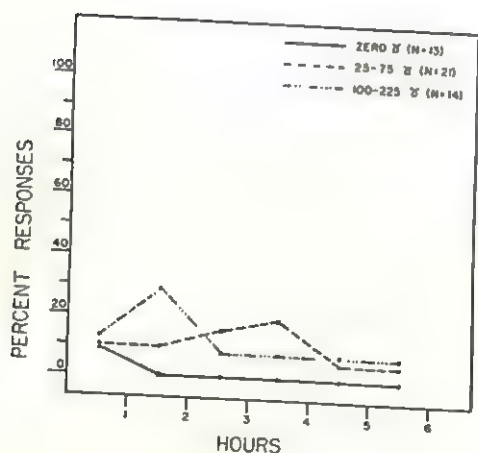
Item 2: "Are you nauseated?" No one with zero dosage reported nausea



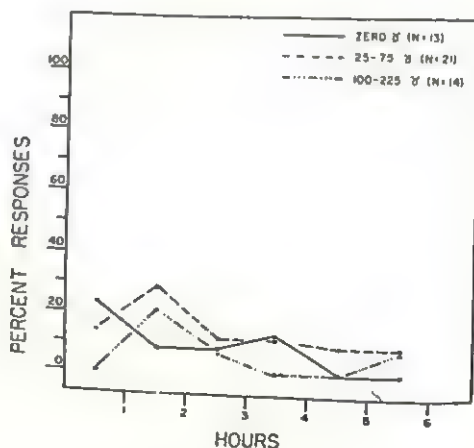
1. DO YOU FEEL ILL IN ANY WAY?



2. ARE YOU NAUSEATED?



3. HAVE YOU A FEELING OF CHOKING?

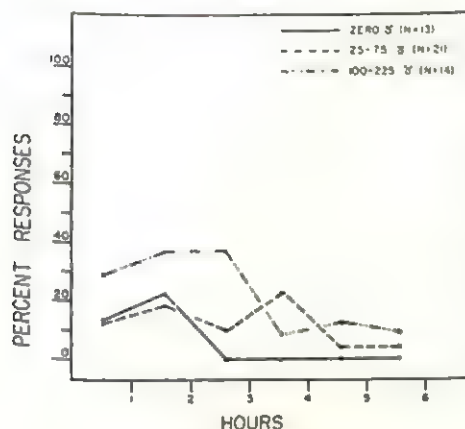


4. IS SALIVATION INCREASED?

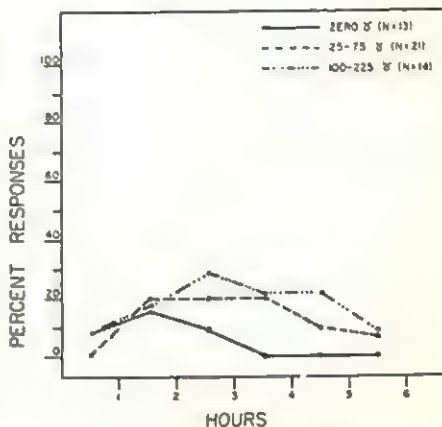
FIGURE 1  
HOURLY RESPONSE TO EACH OF 47 QUESTIONS

until 3½ hours after the placebo, when 15 per cent reported this symptom. For the low dosage level, 25-75 micrograms, the peak number of responses occurred at 2½ hours when 29 per cent of the subjects reported feeling nauseated. Forty-three per cent of the subjects receiving 100-225 micrograms of the drug, reported feeling nauseated 1½ hours later.

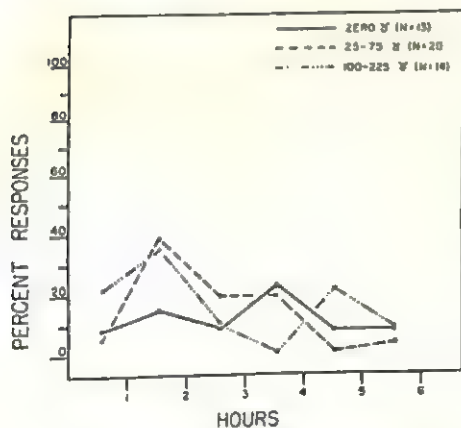
*Item 3: "Have you a feeling of choking?"* There was only one positive response (or 8 per cent) for the first interval for those receiving no LSD-25. Those receiving the low dosage gave the maximum number of responses after 3½ hours when 19 per cent of the subjects gave positive responses. For



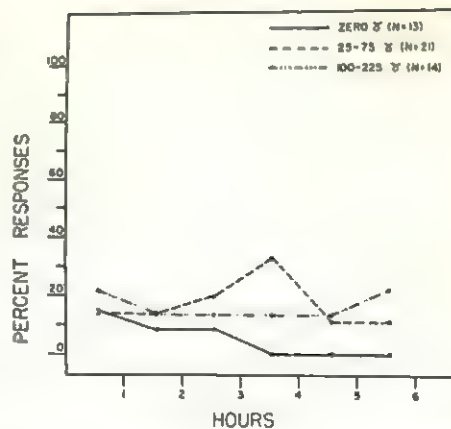
5. IS SALIVATION DECREASED?



6. IS YOUR APPETITE INCREASED?



7. IS YOUR APPETITE DECREASED?



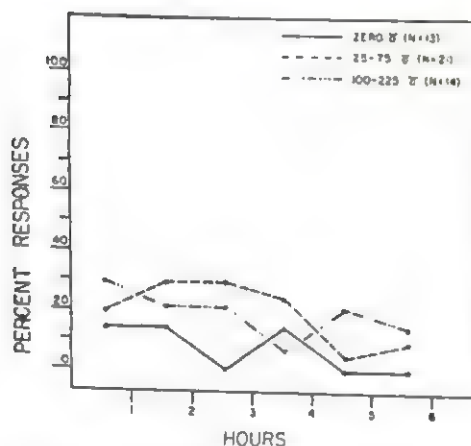
8. DO YOU HAVE A "DRY" TASTE IN YOUR MOUTH?

FIGURE 1 (continued)

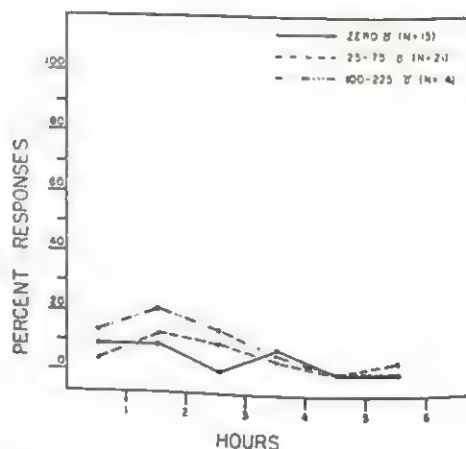
the high dosage the peak was again reached in the second interval (1½ hours) and the percentage giving positive responses was 29. The high dosage of LSD-25 again showed greater and more rapid positive response.

*Item 4: "Is salivation increased?"* Among those receiving no LSD-25, 23 per cent reported that their salivation was increased ½ hour after the placebo. About 29 per cent of the subjects receiving the low dosage reported that their salivation was increased in the second interval. This was also the peak hour for the 100-225 microgram group, when about 21 per cent reported that their salivation was increased.

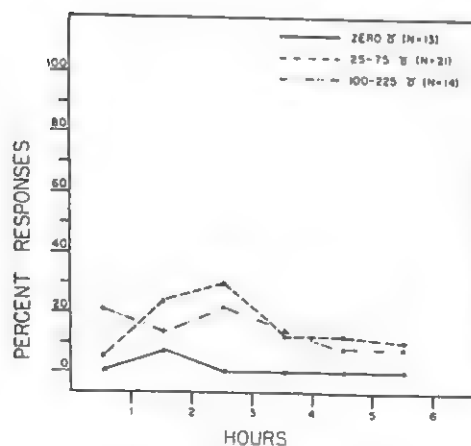
*Item 5: "Is salivation decreased?"* Here the peak was reached at 1½



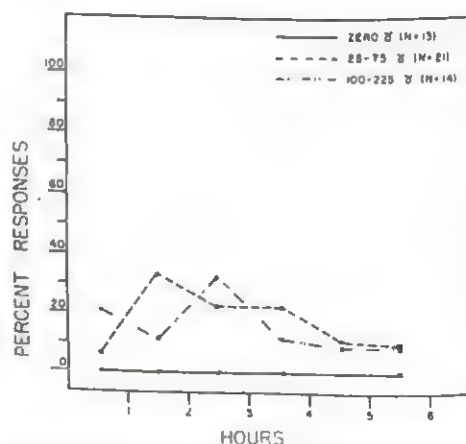
9. DO YOU HAVE A FUNNY TASTE IN YOUR MOUTH?



10. DO YOU HAVE A BITTER TASTE IN YOUR MOUTH?



11. ARE YOUR LIPS NUMB?



12. ARE YOUR LIPS DRAWN BACK AS IF YOU WERE SMILING?

FIGURE 1 (continued)



hours for those receiving no LSD-25, when 23 per cent reported that their salivation was decreased. For those receiving a low dosage the peak was reached 3½ hours after the drug when 24 per cent reported that their salivation was decreased. About 36 per cent receiving the high dosage reported that their salivation was decreased in both the second and third intervals.

*Item 6: "Is your appetite increased?"* It should be noted that the subjects were fed about one-half hour after they took the drug, or just before the first question period. However, they were always instructed to give a positive response only if hunger was increased abnormally. Fifteen per cent of the subjects given no drug reported that their appetites were increased

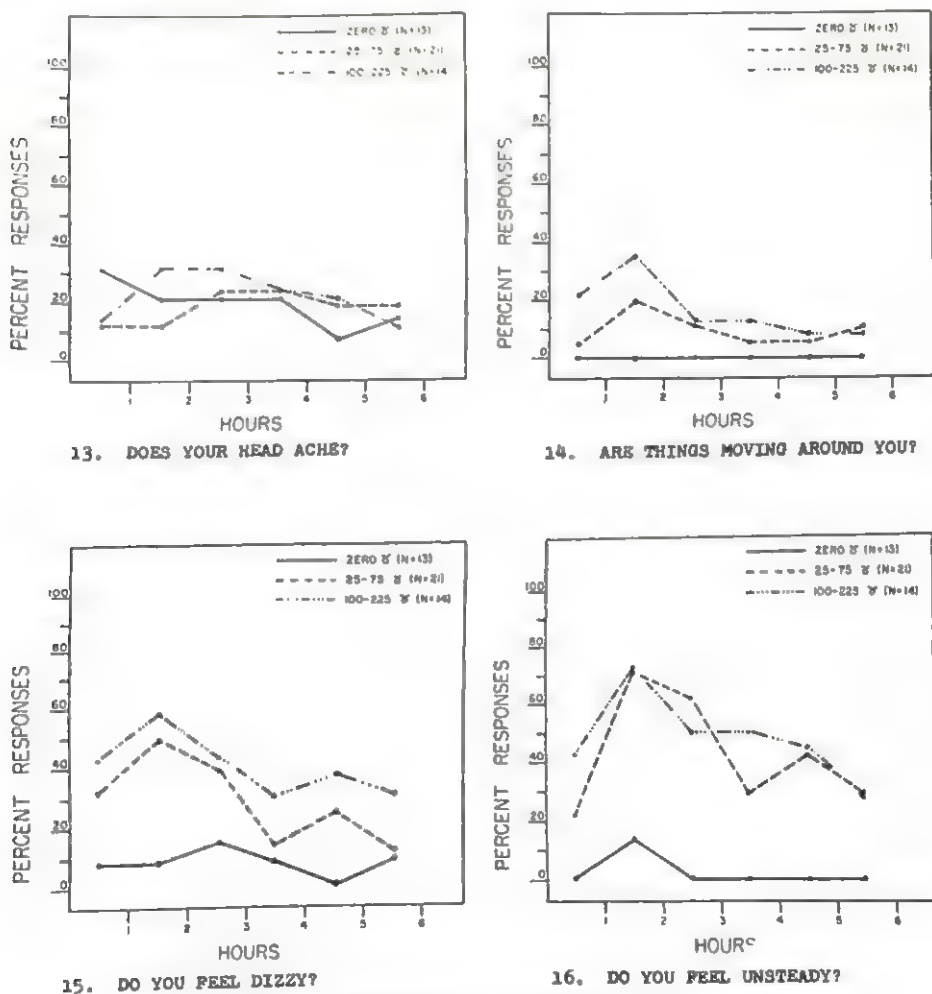
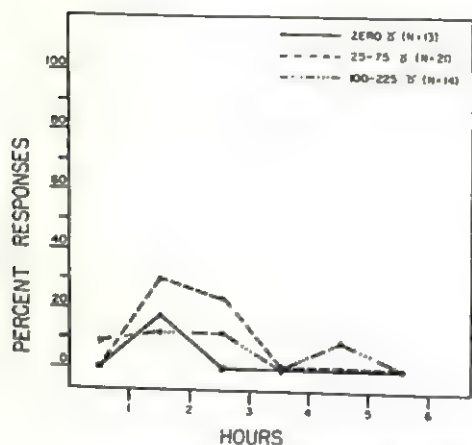


FIGURE 1 (continued)

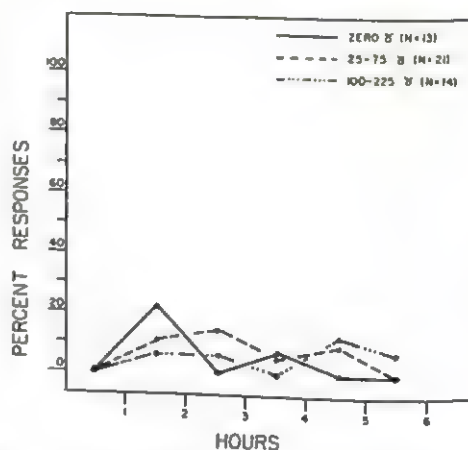
in the second period and about 19 per cent of the subjects receiving a low dose reported an increased appetite from 1½ to 3½ hours after the drug. Twenty-nine per cent of the subjects receiving the high dose reported that their appetites were increased in the third period, the peak for the large dose.

*Item 7: "Is your appetite decreased?"* About 23 per cent of the subjects receiving the placebo reported that appetite decreased after 3½ hours. At 1½ hours, nearly 38 per cent of the subjects receiving the low dose reported a decrease in appetite. Approximately 36 per cent of the subjects receiving 100-225 micrograms of the drug reported a decrease 1½ hours later.

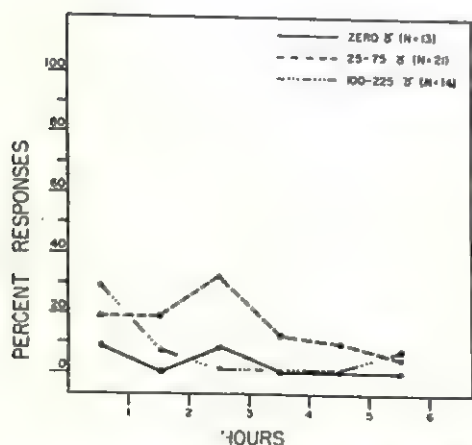
*Item 8: "Do you have a dry taste in your mouth?"* About 15 per cent



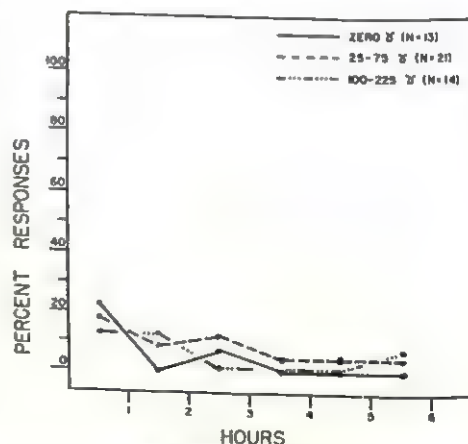
17. IS THERE DIFFICULTY IN BREATHING?



18. DO YOU PASS MORE URINE THAN USUAL?



19. ARE YOU AWARE OF YOUR HEART BEAT?



20. IS YOUR HEART BEAT FASTER THAN USUAL?

FIGURE 1 (continued)

of the persons receiving zero dosage reported in the first interval that they had a dry taste in their mouth. This percentage declined to zero by the fourth hour. On the other hand, the fourth hour was the peak hour for the low dosage group, when approximately 33 per cent of the group responded positively. The high dosage group never yielded a peak hour, and only about 21 per cent of the group reported a dry taste in the first and sixth hours.

Item 9: "Do you have a funny taste in your mouth?" Fifteen per cent of those receiving no LSD-25 reported that they had a funny taste in their mouth, whereas nearly 29 per cent of those receiving the low dose reported a funny taste at 1½ and 2½ hours. Fewer of those receiving the high dosage

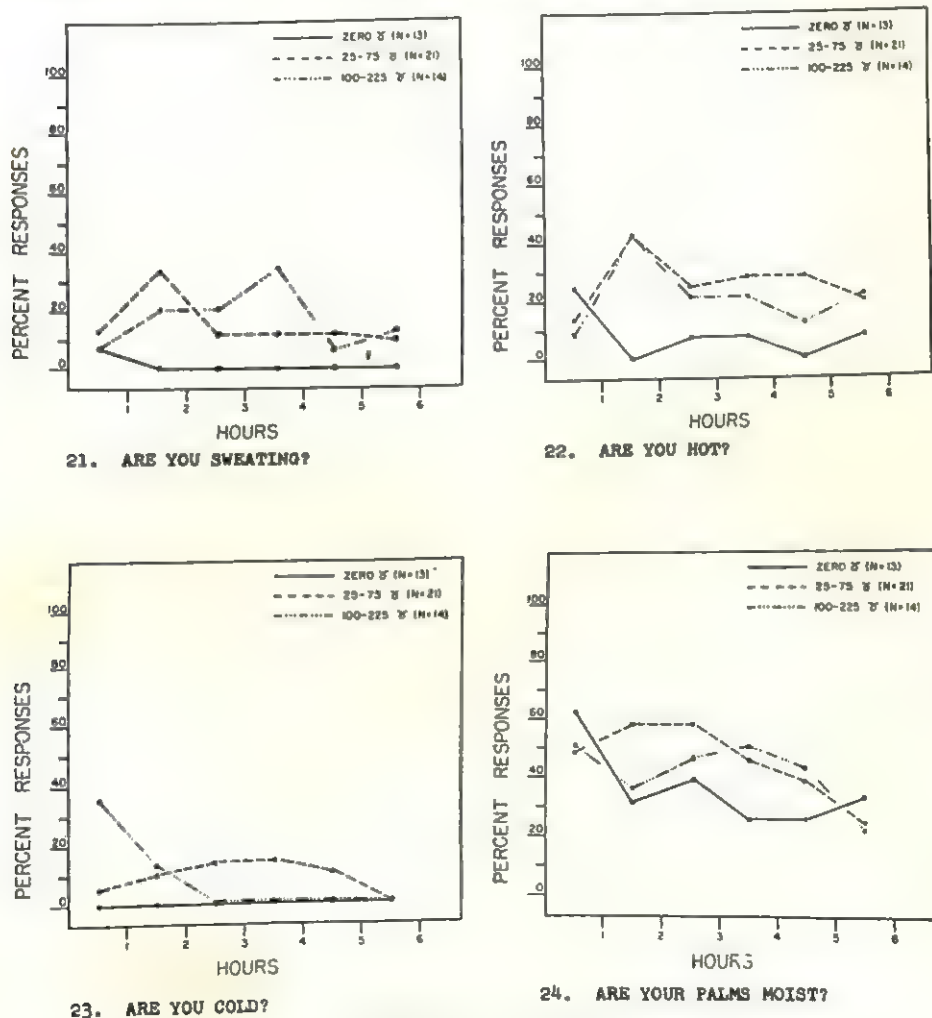


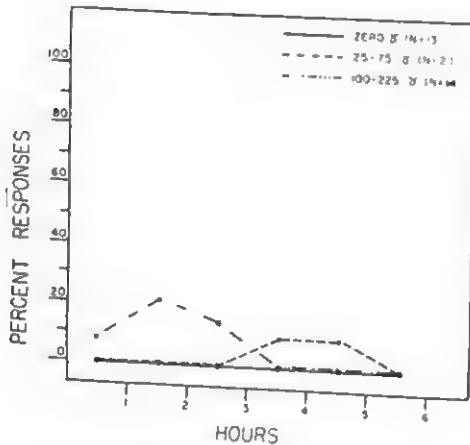
FIGURE 1 (continued)



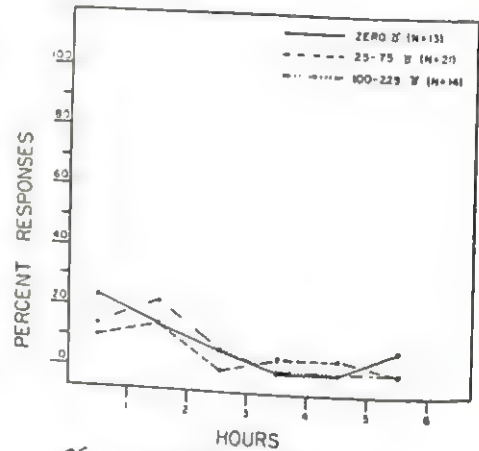
tended to report a funny taste in their mouth, in the same way that fewer tended to report a dry taste in their mouth. The peak with high dosage was again in the first half hour, when a little less than 29 per cent reported a funny taste, with the percentage declining as time progressed.

*Item 10: "Is it a bitter taste?"* Few in any of the groups responded positively to this question. Eight per cent of those receiving no LSD-25 responded positively and 14 per cent of those receiving a low dose responded positively 1½ hours later. The second period was the peak hour for the high dose when 21 per cent responded positively.

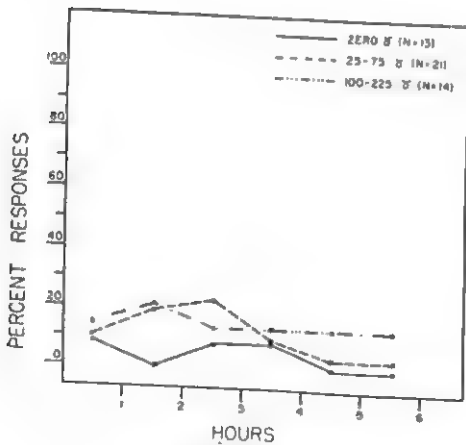
*Item 11: "Are your lips numb?"* Only 8 per cent of those receiving



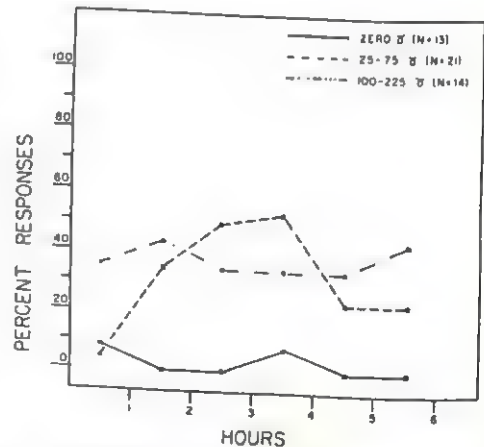
25. ARE YOUR PALMS DRY?



26. ARE YOUR PALMS COLD?



27. IS YOUR SKIN SENSITIVE?

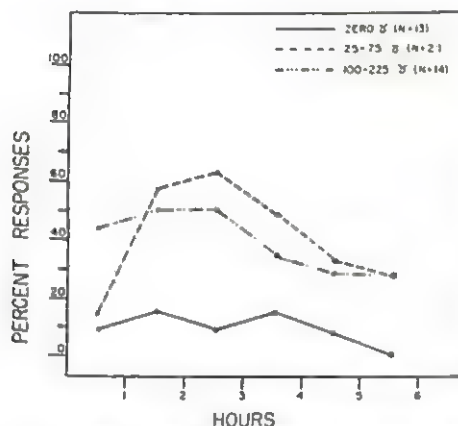


28. DO YOU HAVE FUNNY FEELINGS ON YOUR SKIN?

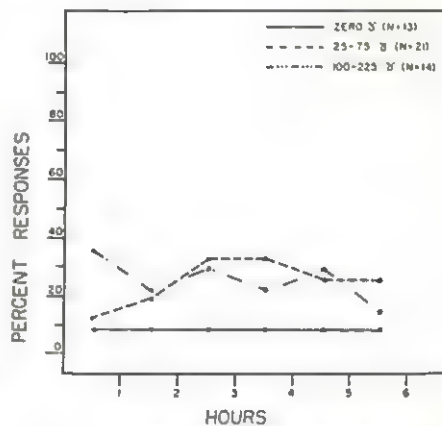
FIGURE 1 (continued)

zero dosage reported that lips were numb in the second period, with no response in any other period. The low dosage group reached its peak at 2½ hours when 29 per cent gave a positive response. The high dosage group began at the peak with 28 per cent, then gradually declined until the sixth hour when only 14 per cent gave positive responses.

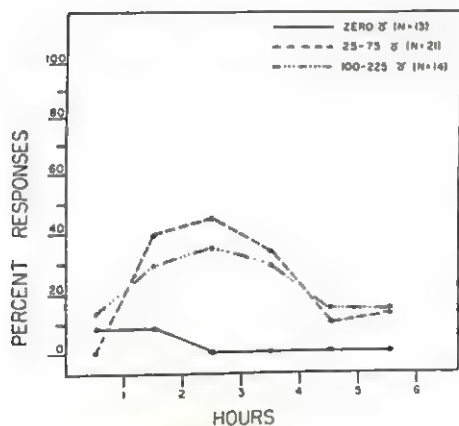
Item 12: "Are your lips drawn back as if smiling?" In this item not one of the subjects in the zero dose group gave a positive response at any hour. Among the subjects receiving the low dose, 33 per cent responded positively at 1½ hours which was the peak hour. For the high dosage group, 21 per cent responded positively in the first and third periods.



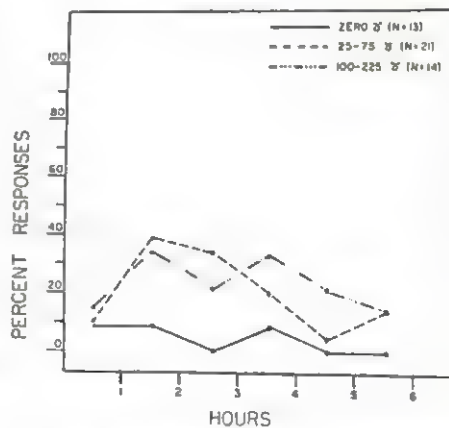
29. DO YOUR HANDS AND FEET FEEL PECULIAR?



30. DO YOUR HANDS AND FEET FEEL HEAVY?



31. DO YOUR HANDS AND FEET FEEL LIGHT?

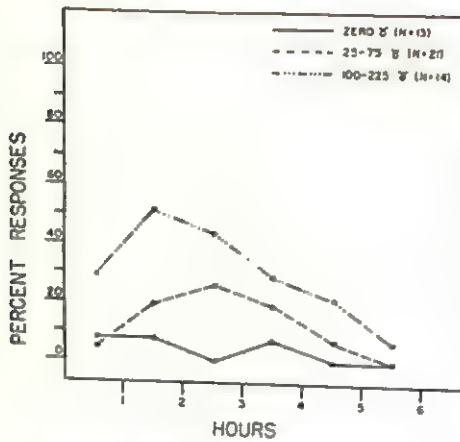


32. IS THERE PRESSURE IN YOUR EARS?

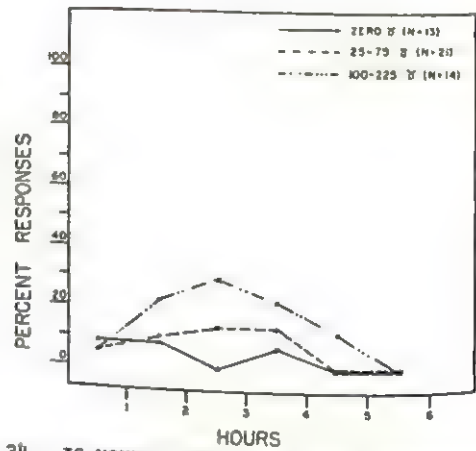
FIGURE 1 (continued)

*Item 13: "Does your head ache?"* A number of the subjects reported headaches whether they got the drug or not. Thirty-one per cent of the subjects who received no LSD-25 reported a headache  $\frac{1}{2}$  hour after the placebo. Twenty-four per cent of the 25-75 microgram group reported a headache at the  $2\frac{1}{2}$  and  $3\frac{1}{2}$  hour intervals. Twenty-nine per cent of the high dosage group reported a headache at  $2\frac{1}{2}$  hours.

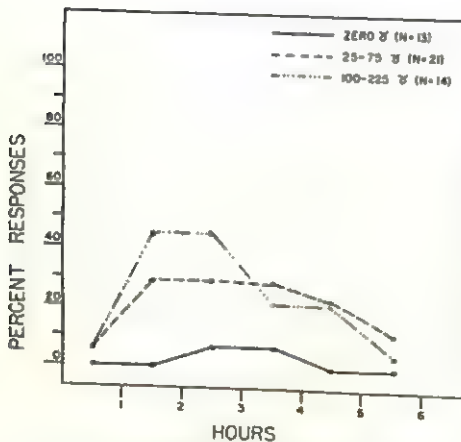
*Item 14: "Are things moving around you?"* None of the subjects who received no LSD-25 reported that things were moving around them. But among those who did receive LSD-25, 19 per cent of the low dosage group



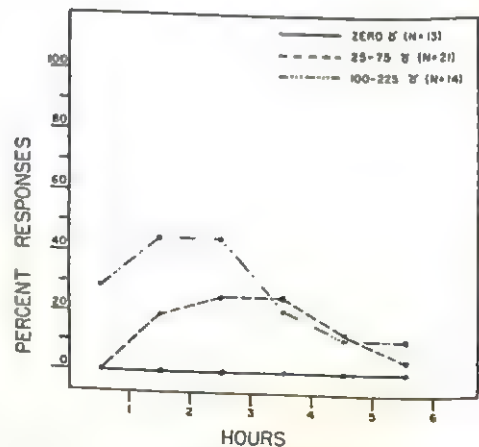
33. IS YOUR HEARING ABNORMAL?



34. IS YOUR HEARING MORE ACUTE THAN USUAL?



35. IS YOUR EYE SIGHT BLURRED?



36. DO YOU HAVE DIFFICULTY IN FOCUSING YOUR VISION?

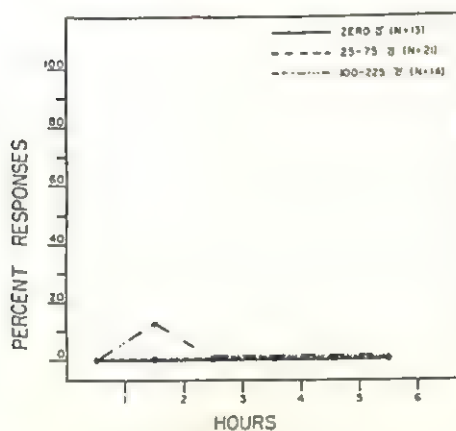
FIGURE 1 (continued)



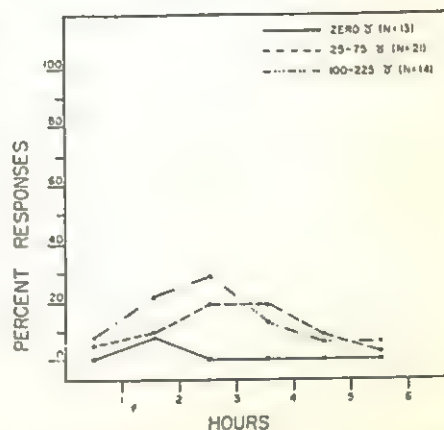
and 36 per cent of the high dosage group answered "yes" 1½ hours after the drug.

*Item 15: "Do you feel dizzy?"* Fifteen per cent of the subjects receiving no LSD-25 felt dizzy in the third period. Of those who received a low dose, 48 per cent reported feeling dizzy in the second period. In the sixth period there were 14 per cent who felt dizzy. Of those receiving the high dose, 57 per cent reported that they felt dizzy in the second interval.

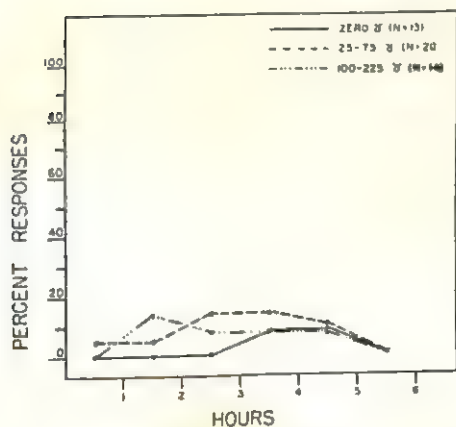
*Item 16: "Do you feel unsteady?"* Fifteen per cent of the subjects receiving the placebo reported that they felt unsteady 1½ hours later, whereas of those receiving the low dosage 71 per cent responded positively at 1½



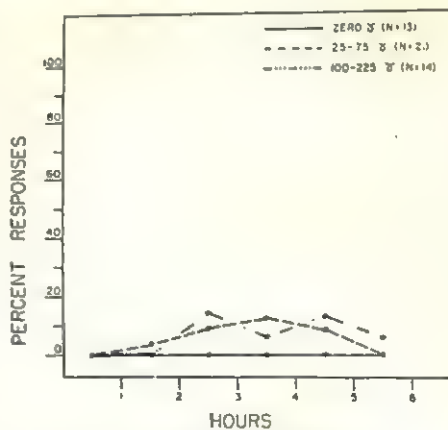
37. DO YOU SEE DOUBLE?



38. ARE SHAPES AND COLORS ALTERED IN ANY WAY?

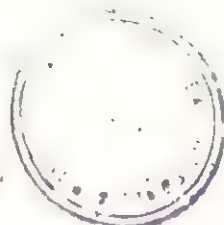
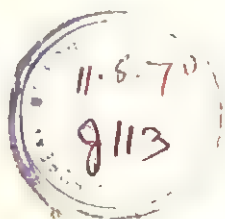


39. DOES LIGHT BOTHER YOU?



40. DO THINGS SEEM TOO CLOSE?

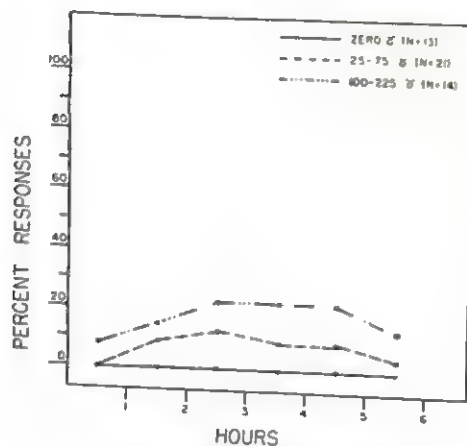
FIGURE 1 (continued)



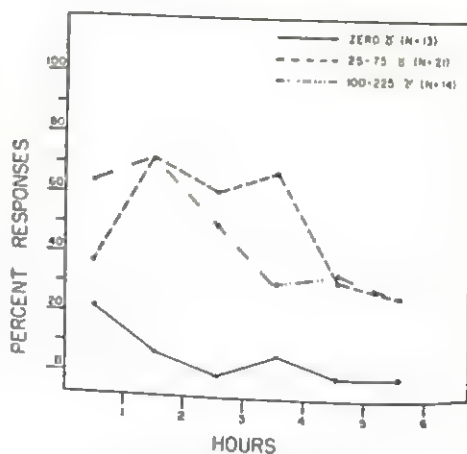
hours. For the high dosage, 71 per cent reported feeling unsteady  $1\frac{1}{2}$  hours after the drug.

*Item 17: "Is there difficulty in breathing?"* Only 8 per cent of the subjects receiving the zero dosage reported difficulty in breathing. In the second period, 29 per cent of the subjects receiving the low dose reported difficulty in breathing, and 14 per cent of the subjects receiving the high dosage reported this feeling. In this item, the highest peak was seen with the low doses.

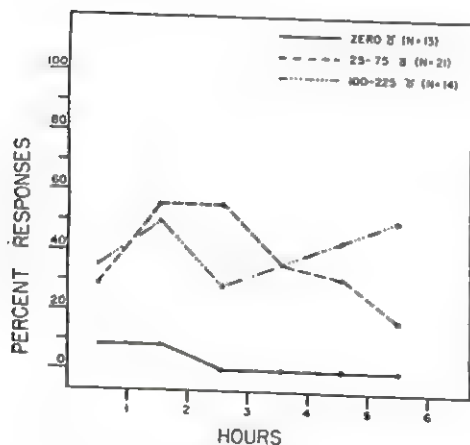
*Item 18: "Do you pass more urine than usual?"* Twenty-three per cent of the subjects who received a placebo said that they were passing more urine



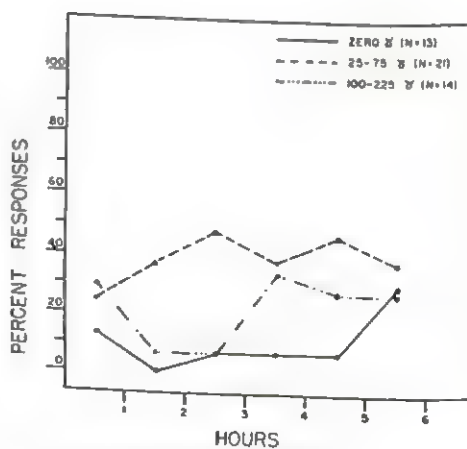
41. DO THINGS SEEM TOO FAR AWAY?



42. DO YOU TREMBLE INSIDE?



43. DO YOU FEEL WEAK?

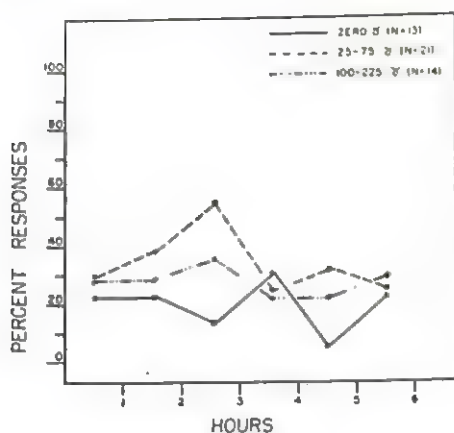


44. DO YOU FEEL FATIGUED?

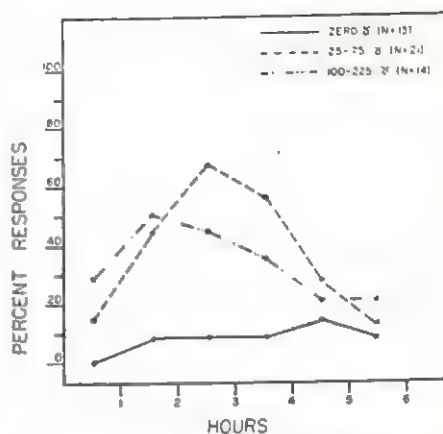
FIGURE 1 (continued)

than usual in the second period, and 14 per cent of those in the low dose group responded positively in the third interval. Seven per cent of the subjects who received the high dosage said that they were passing more urine than usual in the fifth interval.

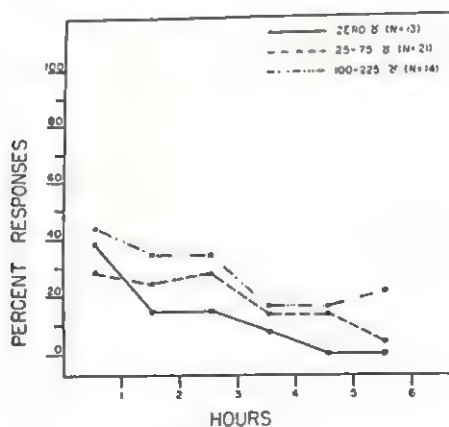
*Item 19: "Are you aware of your heartbeat?"* Only 8 per cent of the subjects who received no LSD-25 responded positively in the first and third periods. Of the subjects who received the low dose, 32 per cent reported that they felt their heartbeat in the third period, which was the peak hour. Of those in the 100-225 microgram group, 29 per cent said that they felt



45. DO YOU FEEL DROWSY?



46. DO YOU FEEL AS IF IN A DREAM?



47. ARE YOU ANXIOUS?

FIGURE 1 (continued)

their heartbeat in the first period. This effect declined rather rapidly, thereafter.

*Item 20: "Is your heartbeat faster than usual?"* Twenty-three per cent of the subjects who received no LSD-25 claimed that their heart rate was faster than usual during the first period. Nineteen per cent of the subjects who received the low dose said their heart rate was faster than usual. Fourteen per cent of the subjects who received the high dose experienced tachycardia at  $\frac{1}{2}$  hour.

*Item 21: "Are you sweating?"* Eight per cent of the subjects who received the placebo said they were sweating in the first period, and never did thereafter. Of the subjects who received a low dose, 33 per cent said that they were sweating in the second period, which was the peak hour. Among the subjects who received the high dose, 36 per cent reported sweating in the fourth period, which was the peak.

*Item 22: "Are you hot?"* Although the room was warm (about 80° F) only 23 per cent of those receiving no LSD-25 said that they were hot,  $\frac{1}{2}$  hour after the placebo. Forty-three per cent of the subjects who received a low dose reported feeling hot  $1\frac{1}{2}$  hours later, and 43 per cent of those in the high dosage group responded positively at  $1\frac{1}{2}$  hours.

*Item 23: "Are you cold?"* Although the room was somewhat warm, some of the subjects who received LSD-25 did say that they felt cold. None of the subjects who did not receive LSD-25 felt cold. In the low dosage group, 14 per cent said that they felt cold in the third and fourth periods. In the high dosage group, 36 per cent said that they felt cold in the first period, and this dropped down rapidly.

*Item 24: "Are your palms moist?"* Many of the subjects who did not receive LSD-25 reported moist palms. In the first  $\frac{1}{2}$  hour, 62 per cent responded positively. The percentage who said that their palms were moist was never less than 24 per cent, and this per cent was not reached until  $3\frac{1}{2}$  hours after the placebo. Of the people who received the low dose, 57 per cent said their palms were moist,  $1\frac{1}{2}$  and  $2\frac{1}{2}$  hours later. Among those in the high dose group, 50 per cent reported moist palms at the first  $\frac{1}{2}$  hour.

*Item 25: "Are your palms dry?"* Few reported dry palms, but those who did always had received LSD-25; no one in the zero dose group did so. Of the people receiving the low dose, 14 per cent said that their palms were dry in the fourth and fifth periods while 21 per cent of the 100-225 microgram group said that their palms were dry in the second period.

*Item 26: "Are your palms cold?"* Of those receiving the placebo, 23 per cent reported that their palms were cold. Among the low dose group, 14 per



cent reported cold palms  $1\frac{1}{2}$  hours after the drug. Twenty-one per cent of those receiving the high dosage reported cold palms  $2\frac{1}{2}$  hours later.

*Item 27: "Is your skin sensitive?"* Eight per cent of the subjects who received zero dosage said that their skin was sensitive. Of those receiving the low dosage 24 per cent said that their skin was sensitive in the third period, and 21 per cent of the people receiving the high dosage responded positively to this item in the second period.

*Item 28: "Do you have funny feelings on your skin?"* Eight per cent responded positively with zero dosage. However, for the group receiving the low dosage, in the fourth period 52 per cent responded positively. In the group receiving the high dosage, 43 per cent responded positively, in the second period. The level remained high for both groups receiving LSD-25.

*Item 29: "Do your hands and feet feel peculiar?"* After the placebo fifteen per cent reported that their hands and feet felt peculiar in the second and the fourth periods. But 62 per cent of those who received a low dose reported that their hands and feet felt peculiar at the third interval; the level was always relatively high in all hours. For the high dose, 50 per cent reported that their hands and feet felt peculiar in the second and third intervals.

*Item 30: "Do they feel heavy?"* Eight per cent of those receiving the placebo answered positively in all hours. Thirty-three per cent of those receiving the low dose answered positively  $2\frac{1}{2}$  and  $3\frac{1}{2}$  hours later, and 36 per cent of those receiving the high dose reported heavy hands and feet  $\frac{1}{2}$  hour after the drug.

*Item 31: "Do they feel light?"* Eight per cent who received zero LSD-25 answered this question positively in the first and second periods, whereas of those receiving the low dose, 43 per cent answered the question positively in the third period. Among the high dosage group, 36 per cent reported a feeling of lightness in the hands and feet during the third period.

*Item 32: "Is there pressure in your ears?"* Only 8 per cent of those receiving zero dosage reported pressure in their ears. Thirty-eight per cent of those receiving the low dosage responded positively,  $1\frac{1}{2}$  hours later. Thirty-six per cent of those receiving the high dosage gave positive answers, with peaks  $1\frac{1}{2}$  and  $3\frac{1}{2}$  hours later.

*Item 33: "Is your hearing abnormal?"* Only 8 per cent of those receiving zero dosage gave positive responses to this item, whereas 24 per cent of those receiving a 25-75 microgram dose of the drug reported abnormal hearing in the third period. Fifty per cent of those receiving the high dosage gave positive answers in the second hour.

*Item 34: "Is your hearing more acute than usual?"* Eight per cent of the subjects in the placebo group responded positively, and 14 per cent of those

receiving a low dosage reported more acute hearing in the third and fourth periods. Twenty-nine per cent of those in the high dose group responded positively in the third period.

*Item 35: "Is your eyesight blurred?"* Eight per cent of the subjects receiving only water said that their eyesight was blurred  $2\frac{1}{2}$  and  $3\frac{1}{2}$  hours later, whereas 29 per cent of those receiving the low dose claimed that their eyesight was blurred from  $1\frac{1}{2}$  to  $3\frac{1}{2}$  hours. Forty-three per cent of those receiving the high dosage said that their eyesight was blurred from  $1\frac{1}{2}$  to  $2\frac{1}{2}$  hours later.

*Item 36: "Do you have difficulty in focusing your vision?"* None of the subjects receiving no LSD-25 responded positively to this item. However, 24 per cent of those receiving the low dosage claimed that they had difficulty focusing in the third and fourth periods. Forty-three per cent of those receiving 100-225 micrograms of LSD-25 had difficulty in focusing their vision in the second and third intervals.

*Item 37: "Do you see double?"* Practically no one saw double. None in the zero dosage and the 25-75 microgram group reported seeing double. Only 14 per cent of those who received a high dosage complained of seeing double, and this was  $1\frac{1}{2}$  hours after the drug.

*Item 38: "Are shapes and colors altered in any way?"* Eighteen per cent of the people receiving zero dose said that the shapes and colors were altered during the second period. Nineteen per cent of those receiving a low dose felt that shapes and colors were altered during the third and fourth periods. Twenty-nine per cent of those receiving the high dosage reported that shapes and colors were altered, primarily in the third interval.

*Item 39: "Does light bother you?"* Only 8 per cent of the placebo group responded positively to this question,  $3\frac{1}{2}$  and  $4\frac{1}{2}$  hours after the water. Fourteen per cent of those receiving a low dose felt that light bothered them  $2\frac{1}{2}$  and  $3\frac{1}{2}$  hours later, while 14 per cent of those in the high dosage group stated that light bothered them at  $1\frac{1}{2}$  hours.

*Item 40: "Do things seem too close?"* No one receiving the zero dosage claimed things seemed too close. Fourteen per cent receiving a low dose complained that things seemed too close in the fourth period. Fourteen per cent of those receiving the high dosage said that things seemed too close in the third and fifth periods.

*Item 41: "Do things seem too far away?"* Those receiving zero dosage always answered negatively. In the third period, 14 per cent of the people receiving a low dosage felt that things seemed too far away. Twenty-one per cent of those receiving a high dosage reported that things seemed too far away in the third, fourth, and fifth periods.

*Item 42: "Do you tremble inside?"* One-half hour after taking the placebo 23 per cent of the people said that they trembled inside, but this rapidly fell off to 8 per cent or less in the succeeding hours. On the other hand, for those who took the drug, 71 per cent of those receiving the low dose and 71 per cent of those who took the high dose complained of inner trembling 1½ hours later.

*Item 43: "Do you feel weak?"* Only 8 per cent of the people receiving no LSD-25 felt weak in the first 1½ hours, whereas 57 per cent of those receiving the low dosage felt weak 1½ and 2½ hours later. Half of those receiving a high dosage felt weak at 1½ and 5½ hours.

*Item 44: "Do you feel fatigued?"* The number of positive responses to this item went up toward the end of the experiment. For those subjects who received zero dosage, it went up as high as 31 per cent in the sixth period. Forty-eight per cent reported it in the third and fifth periods (low dosage). On the other hand, among those receiving 100-225 gamma, the percentage giving positive responses was only 36 in the fourth period.

*Item 45: "Do you feel drowsy?"* Thirty-one per cent receiving a placebo felt drowsy in the fourth interval. Fifty-seven per cent of the subjects receiving the low dosage of LSD-25 felt drowsy in the third period, and 36 per cent of those in the high dosage group felt drowsy in the third period.

*Item 46: "Do you feel as if in a dream?"* Fifteen per cent of the subjects receiving zero dose felt as though they were in a dream 4½ hours later. The subjects receiving a 25-75 microgram dose of LSD-25 felt as if in a dream mainly at 2½ hours, when 67 per cent of them answered "Yes" to this question. Those in the high dosage group reached their peak at 1½ hours, when 50 per cent of them reported experiencing a dream-like state.

*Item 47: "Are you anxious?"* For all subjects, whether or not they received LSD-25, the total number of responses was highest in the first period, and declined thereafter. Thirty-eight per cent of the subjects who received a placebo reported anxiety in the first period. By the fifth period, no one reported it. For those receiving a low dose, 29 per cent responded positively in the first period, declining to 4 per cent in the sixth period. Among those in the high dose group, 43 per cent stated that they were anxious in the first period, and by the fourth and fifth periods, only 14 per cent reported it.

*Average percentage positive responses to 47 questions (Figure 2).* When responses to all items for the zero dose group are combined, there is at first a linear decrease with time in percentage of responses with the exception of a slight peak at 3½ hours. Compare this with the curve for the low dose group. Following an initially low level of responses there is a sharp rise to a maximum at 2½ hours. There is then a gradual decline.

The high dosage curve begins at a relatively high level in the first  $\frac{1}{2}$  hour; it reaches its peak, which is higher than that of the low dose, at  $1\frac{1}{2}$  hours. Gradual decline follows.

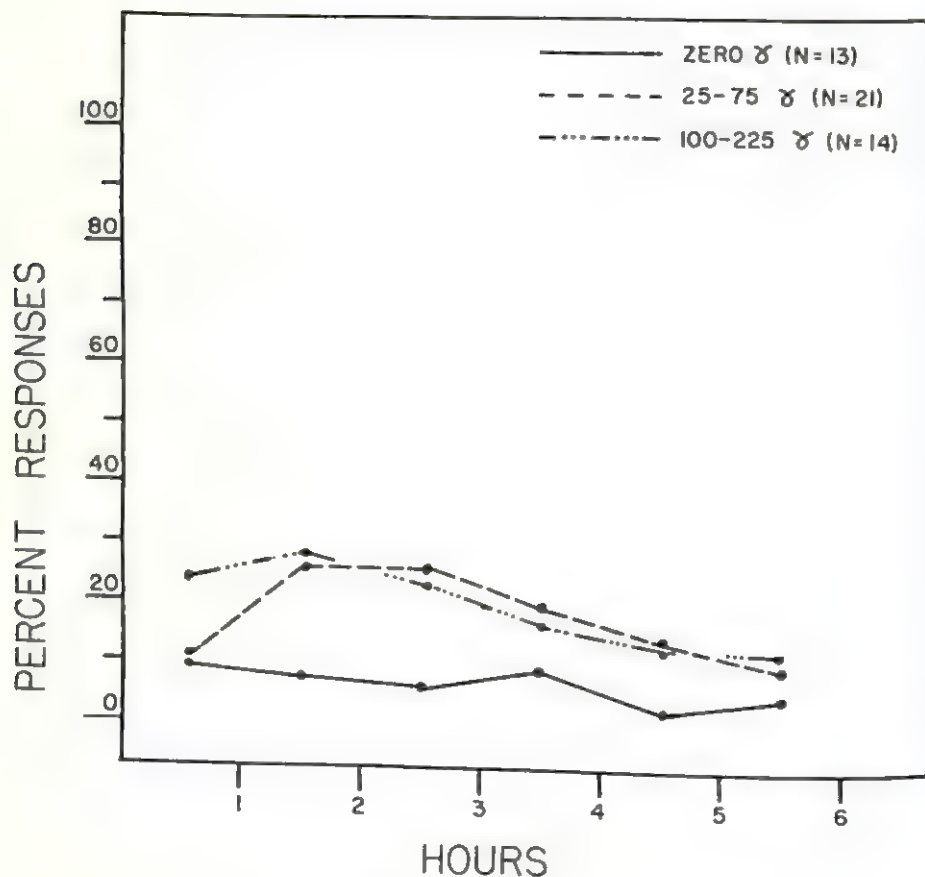


FIGURE 2  
AVERAGE PERCENT POSITIVE RESPONSE TO 47 QUESTIONS

(3). *Discussion.* The symptoms which seem most closely related to the drugs, by virtue of their receiving considerably larger responses by the drug groups than by the placebo group, are the following: things seeming to move around subjects, dizziness, unsteadiness, sweating, feeling hot or cold, paraesthesias, peculiar feeling of limbs, heaviness or lightness of limbs, pressure in ears, blurred eyesight, difficulty in focusing vision, inner trembling, weakness and dream-like feeling.

Some items differentiated between the placebo group and the drug group to a lesser extent. These symptoms are: choking, changes in appetite, dry and funny taste in mouth, numbness of lips, lips drawn back as if smiling, aware-



ness of heartbeat, hyperacusia, alteration of shapes and colors, and objects appearing too far away. Illness, nausea, decreased salivation, and abnormal hearing did not seem to occur frequently until higher doses of the drug were taken.

In addition to symptoms which seem to be drug-produced, other differences appear among the three groups. As a whole, most of the responses by the group taking zero dosage are given  $\frac{1}{2}$  hour after taking the placebo and taper off quickly. Under the low doses of LSD-25, symptoms generally attain a peak between  $1\frac{1}{2}$  and  $2\frac{1}{2}$  hours after the drug. They then decline gradually and are often still reported by the last hour. Higher doses seem to take effect much more rapidly, and the peak is usually reached at  $1\frac{1}{2}$  hours. The decline is much more gradual than under the low dose and the level of response is usually higher throughout.

If the three groups of subjects were correlated, more accurate results would be obtained since individual variability would be controlled. However, these results do give some indication of LSD-25 symptomatology.

One other weakness must be pointed out: in the case of symptoms reported frequently by all three groups there is no way of knowing whether the feelings experienced are subjectively the same or whether the drug actually does produce different qualitative results in areas receiving the same positive responses. The questionnaire is not sufficiently definitive for this, and it is difficult to construct one which does measure qualitative differences in such terms as "Do you feel drowsy?" or "Does your head ache?"

*b. Average percentage response per hour.*

(1). *Statistical procedure.* The total number of plus (or positive) responses made by each group to each question throughout the six-interval day was counted by the IBM 101 Electronic Statistical Machine. The average percentage responding per hour was then determined by means of the following formula:

$$M_{\%} = \frac{\Sigma +}{6N} \times 100 \quad (2)$$

where  $M_{\%}$  equals the average percentage responding positively per hour,  $\Sigma +$  equals the sum of plus responses throughout six intervals, and  $N$  equals the number of subjects.

The groups are the same as those described in the previous section on hourly percentage responses to questions. The number of subjects equals 13, 21, and

14 for the zero, 25-75 micrograms, and 100-225 microgram groups, respectively.

(2). *Results.* It can be seen from Table 2 that in 26 of the 47 questions, increasing percentages of responses are made under increasing dosages of the drug. These questions are: (1) "Do you feel ill in any way?", (2) "Are you nauseated?", (3) "Have you a feeling of choking?", (5) "Is salivation decreased?", (6) "Is your appetite increased?", (7) "Is your appetite decreased?", (10) "Is it (the funny taste) a bitter taste?", (11) "Are your lips numb?", (14) "Are things moving around you?", (15) "Do you feel dizzy?", (16) "Or (do you feel) unsteady?", (21) "Are you sweating?", (25) "Are your palms dry?", (27) "Is your skin sensitive?", (28) "Do you have funny feelings on your skin?", (30) "Do they (your hands and feet) feel heavy?", (31) "Or (do they feel) light?", (32) "Is there pressure in your ears?", (34) "Is it (your hearing) more acute than usual?", (35) "Is your eyesight blurred?", (36) "Do you have difficulty in focusing your vision?", (38) "Are shapes and colors altered in any way?", (41) "Or (do things seem) too far away?", (42) "Do you tremble inside?", (43) "Do you feel weak?", and (47) "Are you anxious?".

In 17 questions the highest percentage of response occurs at intermediate doses of LSD-25. Among these are: (4) "Is salivation increased?", (8) "Do you have a dry taste in your mouth?", (12) "Are your lips drawn back as if you were smiling?", (13) "Does your head ache?", (17) "Is there difficulty in breathing?", (18) "Do you pass more urine than usual?", (19) "Are you aware of your heartbeat?", (20) "Is it faster than usual?", (22) "Are you hot?", (23) "Or (are you) cold?", (24) "Are your palms moist?", (29) "Do your hands and feet feel peculiar?", (33) "Is your hearing abnormal?", (39) "Does light bother you?", (44) "Or (do you feel) fatigued?", (45) "Do you feel drowsy?", and (46) "Do you feel as if in a dream?".

Some of the differences between percentages of the different groups responding are relatively small and others are large. Whether these differences are real differences or whether they are a result of chance cannot be determined by inspection. Results of statistical treatment of the differences are shown in Tables 8-10.

Equal percentages of positive responses were made by the two drug groups to the questions: (9) "Do you have a funny taste in your mouth?", (40) "Do things seem too close?", and (26) "Or (are your palms) cold?", with a larger per cent responding under zero than under the drug in the case of the latter question only. (37) "Do you see double?", received equal response by the zero group and the 25-75 gamma group, with a slightly larger per cent responding under the high dose.

Under zero dosage, (24) "Are your palms moist?", received approximately one-third positive responses per hour, while (45) "Are you drowsy?" was positively responded to by about one-fourth of the group. Twenty per cent reported a headache, 13 per cent anxiety and 11 per cent stated they felt decreased appetite and fatigue. Less than 10 per cent of the subjects responded positively to all the other items. These findings are more readily seen in Table 3

There were eight items to which none of the subjects receiving no LSD-25 responded positively (see Table 3). These were: (12) "Are your lips drawn back as if you were smiling?", (14) "Are things moving around you?", (23) "Are you cold?", (25) "Are your palms dry?", (36) "Do you have difficulty in focusing your vision?", (37) "Do you see double?", (40) "Do things seem too close?", and (41) "Do things seem too far away?".

#### Table 4

While only one item received over 30 per cent average hourly response by the zero dose group, nine were responded to at this frequency by subjects under 25-75 micrograms of the drug (see Table 4). These included: (24) "Are your palms moist?", (16) "Do you feel unsteady?", (29) "Do your hands and feet feel peculiar?", (42) "Do you tremble inside?", (43) "Do you feel weak?", (44) "Do you feel fatigued?", (46) "Do you feel as if in a dream?", (45) "Do you feel drowsy?", and (28) "Do you have funny feelings on your skin?".

Thirteen symptoms were reported by no more than 10 per cent of the low dose group (see Table 4): faster heartbeat, things moving around, difficulty in breathing, coldness, photophobia, things seeming too far away, bitter taste in mouth, polyuria, cold palms, hyperacusia, things seeming too close, dry palms, and double vision. Items are listed in order of decreasing frequency and only the last symptom had no response at all under this dose of drug.

When subjects received 100-225 micrograms of LSD-25 the item to which they responded positively most frequently (see Table 5) was (16) "Do you feel unsteady?" Nearly half of the subjects in this dosage range gave positive responses to this item. Almost as many responded positively to (42) "Do you tremble inside?" At this dosage there were also nine questions to which more than 30 per cent of the subjects responded positively. Of these nine questions to which more than 30 per cent of the group receiving the high dosage responded positively, seven were also the same questions that more than 30 per cent of the group receiving the low dosage responded positively. The only two items which were different, and which therefore had moved up in percentage of response were: (15) "Do you feel dizzy?", to which 40 per cent of the subjects receiving the high dosage responded positively, (whereas only 29

TABLE 3

AVERAGE PERCENT RESPONDING, POSITIVELY UNDER ZERO LSD-25  
 (Questions are ranked from highest percent (rank 1) of subjects responding  
 positively to lowest percent.  $N = 13$ )

Rank	Percent responding	Question number	Question
1	35	24	Are your palms moist?
2	24	45	Do you feel drowsy?
3	20	13	Does your head ache?
4	13	47	Are you anxious?
5.5	11	7	Is your appetite decreased?
5.5	11	44	Do you feel fatigued?
8	9	4	Is salivation increased?
8	9	26	Are your palms cold?
8	9	29	Do your hands and feet feel peculiar?
10.5	8	1	Do you feel ill in any way?
10.5	8	9	Do you have a funny taste in your mouth?
13	7	15	Do you feel dizzy?
13	7	22	Are you hot?
13	7	30	Do your hands and feet feel heavy?
16	6	5	Is salivation increased?
16	6	42	Do you tremble inside?
16	6	46	Do you feel as if in a dream?
19.5	5	6	Is your appetite increased?
19.5	5	8	Do you have a dry taste in your mouth?
19.5	5	18	Do you pass more urine than usual?
19.5	5	20	Is your heart beat faster than usual?
24	4	10	Do you have a bitter taste in your mouth?
24	4	27	Is your skin sensitive?
24	4	32	Is there pressure in your ears?
24	4	33	Is your hearing abnormal?
24	4	34	Is it more acute than usual?
28	3	2	Are you nauseated?
28	3	16	Do you feel unsteady?
28	3	17	Is there difficulty in breathing?
32.5	2	19	Are you aware of your heart beat?
32.5	2	28	Do you have funny feelings on your skin?
32.5	2	31	Do your hands and feet feel light?
32.5	2	35	Is your eyesight blurred?
32.5	2	39	Does light bother you?
32.5	2	43	Do you feel weak?
37.5	1	3	Have you a feeling of choking?
37.5	1	11	Are your lips numb?
37.5	1	21	Are you sweating?
37.5	1	38	Are shapes and colors altered in any way?
43.5	0	12	Are your lips drawn back as if you were smiling?
43.5	0	14	Are things moving around you?
43.5	0	23	Are you cold?
43.5	0	25	Are your palms dry?
43.5	0	36	Do you have difficulty in focusing your vision?
43.5	0	37	Do you see double?
43.5	0	40	Do things seem too close?
43.5	0	41	Do things seem too far away?



TABLE 4  
AVERAGE PERCENT RESPONDING POSITIVELY UNDER 25-75 GAMMA LSD-25  
(Questions are ranked from highest percent (rank 1) of subjects responding positively to lowest percent. N = 21)

Rank	Percent of subjects responding positively	Question number	Question
1	45	24	Are your palms moist?
2	43	16	Do you feel unsteady?
3	41	29	Do your hands and feet feel peculiar?
4	40	42	Do you tremble inside?
5.5	39	43	Do you feel weak?
5.5	39	44	Do you feel fatigued?
7	37	46	Do you feel as if in a dream?
8	34	45	Do you feel drowsy?
9	31	28	Do you have funny feelings on your skin?
10	29	15	Do you feel dizzy?
11	27	13	Does your head ache?
12	26	22	Are you hot?
13	25	30	Do your hands and feet feel heavy?
14	23	31	Do your hands and feet feel light?
15.5	22	35	Is your eyesight blurred?
15.5	22	47	Are you anxious?
17	20	32	Is there pressure in your ears?
18.5	19	9	Do you have a funny taste in your mouth?
18.5	19	12	Are your lips drawn back as if you were smiling?
21	17	8	Do you have a "dry" taste in your mouth?
21	17	19	Are you aware of your heart beat?
21	17	21	Are you sweating?
23	16	11	Are your lips numb?
24.5	15	2	Are you nauseated?
24.5	15	4	Is salivation increased?
26.5	14	7	Is your appetite decreased?
26.5	14	36	Do you have difficulty in focusing your vision?
28.5	12	5	Is salivation decreased?
28.5	12	33	Is your hearing abnormal?
32	11	1	Do you feel ill in any way?
32	11	3	Have you a feeling of choking?
32	11	6	Is your appetite increased?
32	11	27	Is your skin sensitive?
32	11	38	Are shapes and colors altered in any way?
35	10	20	Is your heart beat faster than usual?
37	9	14	Are things moving around you?
37	9	17	Is there difficulty in breathing?
37	9	23	Are you cold?
39.5	8	39	Does light bother you?
39.5	8	41	Do things seem too far away?
43	7	10	Do you have a bitter taste in your mouth?
43	7	18	Do you pass more urine than usual?
43	7	26	Are your palms cold?
43	7	34	Is your hearing more acute than usual?
43	7	40	Do things seem too close?
46	3	25	Are your palms dry?
47	0	37	Do you see double?

TABLE 5

PERCENT RESPONDING POSITIVELY UNDER 100-225 GAMMA LSD-25

(Questions ranked from highest percent (rank 1) of subjects responding to lowest percent.  $N = 14$ .)

Rank	Percent responding	Question number	Question
1	48	16	Do you feel unsteady?
2	47	42	Do you tremble inside?
3.5	41	24	Are your palms moist?
3.5	41	43	Do you feel weak?
5.5	40	15	Do you feel dizzy?
5.5	40	29	Do your hands and feet feel peculiar?
7	38	28	Do you have funny feelings on your skin?
8	33	46	Do you feel as if in a dream?
9	30	30	Do your hands and feet feel heavy?
10	28	45	Do you feel drowsy?
11.5	27	36	Do you have difficulty in focusing your vision?
11.5	27	47	Are you anxious?
13	26	2	Are you nauseated?
15	24	1	Do you feel ill in any way?
15	24	32	Is there pressure in your ears?
15	24	35	Is your eyesight blurred?
17.5	23	31	Do your hands and feet feel light?
17.5	23	44	Do you feel fatigued?
19	22	5	Is salivation decreased?
20.5	21	11	Are your lips numb?
20.5	21	22	Are you hot?
22	20	13	Does your head ache?
23.5	19	7	Is your appetite decreased?
23.5	19	9	Do you have a funny taste in your mouth?
25.5	18	21	Are you sweating?
25.5	18	27	Is your skin sensitive?
27.5	17	6	Is your appetite increased?
27.5	17	14	Are things moving around you?
29	16	8	Do you have a dry taste in your mouth?
30	15	34	Is your hearing more acute than usual?
31.5	14	12	Are your lips drawn back as if you were smiling?
31.5	14	38	Are shapes and colors altered in any way?
33	13	41	Do things seem too far away?
34	12	3	Have you a feeling of choking?
35	9	10	Do you have a bitter taste in your mouth?
36.5	8	20	Is your heart beat faster than usual?
36.5	8	23	Are you cold?
40	7	17	Is there difficulty in breathing?
40	7	19	Are you aware of your heart beat?
40	7	25	Are your palms dry?
40	7	26	Are your palms cold?
40	7	40	Do things seem too close?
44	6	4	Is salivation increased?
44	6	18	Do you pass more urine than usual?
44	6	39	Does light bother you?
46	3	33	Is your hearing abnormal?
47	2	37	Do you see double?

per cent of those receiving the low dosage did so), and (30) "Do your hands and feet feel heavy?", to which 30 per cent of the subjects responded positively with the high dosage, and 25 per cent of the subjects receiving the low dosage responded positively. Fatigue and drowsiness were less important to the high dosage group and fewer than 20 per cent reported these symptoms.

There were 13 items which received positive responses by no more than 10 per cent of the group (see Table 5). Of these 13 items, there were 9 items in common with the low dose group: (37) "Do you see double?", (39) "Does light bother you?", (18) "Do you pass more urine than usual?", (40) "Do things seem too close?", (26) "Are your palms cold?", (25) "Are your palms dry?", (17) "Is there difficulty in breathing?", (23) "Are you cold?", (10) "Do you have a bitter taste in your mouth?" These items were rarely responded to by subjects receiving LSD-25, regardless of dose.

While none of the questions received zero response from the 100-225 microgram group, the least important question was (37) "Do you see double?" to which an average of 2 per cent per hour responded positively.

(3). *Discussion.* Some of the most interesting data emanate from the zero dosage group. It is interesting that even in this group of individuals who received no LSD-25, but only a placebo consisting of a cup of water, 35 per cent of the responses to "Are your palms moist?" were positive. Several reasons for this large number of positive responses are suggested. First, it may be that individuals normally have moist palms but are not made aware of it until they are asked a direct question. Secondly, the subjects may have been anxious in the situation, and it is well known that anxiety causes moist palms due to perspiration. Although a good many of the sessions were held in a very warm room, the high incidence of moist palms cannot be attributed to this since only one per cent of the subjects responded that they were sweating, when they did not receive LSD-25. The group receiving the drug gave similar high percentage responses to this question and it appears that this symptom is independent of the drug action, but results from our test situation.

The next question to receive a large number of responses by the zero dose group was "Do you feel drowsy?". Twenty-four per cent of the responses were positive. Drowsiness is one of the symptoms which is most amenable to suggestion. The high incidence of drowsiness may also be due to the fact that subjects were situated in a very warm and restful atmosphere, and in addition they were probably fatigued by the rigorous series of tests to which they were submitted. Thirty-four per cent and 28 per cent of subjects in the low and high dose groups, respectively, also gave positive responses to this item showing that this question is not especially discriminatory. It is very likely that the reasons suggested for the high response under zero dosage are applicable to the

drug groups, as well. It is interesting that the highest percentage response occurred with the 25-75 microgram group. The significance of this observation is questionable, however, since the three dose groups were only partially correlated, and the samples were small. Small differences in percentages cannot be given too much significance because of the individual variability present.

The next most common symptom in zero dosage was headache. One out of five responses, on an average, was positive to this. Why there are so many positive responses is difficult to say. Interestingly enough, under 100-225 micrograms, there were 20 per cent giving positive responses to this item, and under 25-75 micrograms, 27 per cent gave positive responses. Regardless of whether the subject took LSD-25 or what the dosage was, about one-fifth of the subjects reported headache. This is not a discriminating question.

Question (47) "Are you anxious?" received positive responses by 13 per cent of the zero dose group; this was the fourth most frequently reported symptom with this dose. Larger percentages of the subjects taking the drug responded to this question indicating that the item probably significantly discriminates. Differences were not great, however, and reliability of the differences would have to be tested statistically. The stress situation of taking a new drug, and not knowing the effects it would have, seem to contribute to the positive responses made in this case.

Decreased appetite reported by zero dose subjects seems related to anxiety. Slightly greater percentages of the group under the effects of LSD-25 responded positively to this item, perhaps relating to the increased anxiety reported by these groups as compared to the placebo group.

Fatigue which was reported might be a result of the general situation of being tested all day. It is strange that the low dose group responded positively to a much greater extent than either of the other groups; it should be remembered, however, that subjects were not identical in each group.

Among the symptoms reported infrequently or never by the zero group, those which seem to have some relationship to LSD-25 were: difficulty in focusing vision, blurred eyesight, inner trembling, dream-like feeling, weakness, pressure in ears, funny feelings on the skin, peculiar feeling of limbs, heaviness and lightness of limbs, nausea, sweating, and feeling hot. Whether or not the differences in percentage response to these questions are significant will be shown in a later discussion.

The average hourly percentage response as a means of comparing the three groups is both advantageous and disadvantageous. Its advantage lies in the fact that there is one representative statistic for each group which can be compared with greater ease than six numbers representing six hours. There is a disadvantage, however, because a great deal of valuable information is



lost. If the reader knows only the average per cent response to each item there is a tendency to assume that shape of the time curve is similar for the three groups. That this is not so was shown in the graphs of each question. Frequently the maximum response under a placebo occurred in the first half-hour and then leveled off while for the drug group the peak usually appeared  $1\frac{1}{2}$  to  $2\frac{1}{2}$  hours after ingestion of the LSD-25. Consequently, some of the questions receiving a relatively low per cent response by zero dose subjects received initially high responses, but rapid leveling off reduced the the average hourly response considerably. It is still of value to compare the average responses, but the nature of the difference is more clearly understood through study of the responses at each hour.

On the basis of the average hourly percentage responses, it would appear that the characteristic feature of items which elicit positive responses under LSD-25, is that they are all related to autonomic functions, especially to autonomic functions which may be said to be elicited by anxiety. For example, the top items of the 25-75 microgram LSD-25 dosage involved such symptoms as moist palms, unsteadiness, peculiar feelings of hands and feet, inner trembling, and weakness. Similarly, with the high dosage of LSD-25 the symptoms which were most frequently elicited were unsteadiness, inner trembling, moist palms, weakness, dizziness, and peculiar feelings of hands and feet. By contrast, the most frequent symptoms of the zero dosage group were (aside from the moist palms) drowsiness, headache, anxiety, decreased appetite, and fatigue. These are mostly depressive symptoms as compared with the orthosympathetic symptoms elicited by LSD-25. In a sense, the symptoms which appear under zero dosage might be related to those that are brought forth by cholinergic drugs (or histamines), whereas the symptoms elicited by LSD-25 mainly resemble those that are elicited by adrenergic drugs.

*c. Ranks of questions.*

(1). *Statistical procedure.* Questions were ranked within each of the three dose groups (zero, 25-75 micrograms and 100-225 micrograms of LSD-25), according to the average percentage responding per hour as indicated in Table 3. Rank "1" was delegated to that question having the greatest percentage response, rank "2" to that having the second highest percentage response, etc.

The relationship among rank positions of questions in the three groups of subjects was determined by applying equation (3) to the ranks comparing the three groups taken two at a time, (zero with 25-75 micrograms, zero with 100-225 micrograms, and 25-75 micrograms with 100-225 micrograms),

$$\rho = 1 - \frac{6\sum D^2}{N(N^2 - 1)} \quad (3)$$

where  $\rho$  equals the rank difference correlation coefficient,  $D$  equals the difference between the ranks of a question in two different dose groups,  $\sum D^2$  equals the sum of the squared rank differences for all 47 items in the groups compared, and  $N$  equals the number of questions, 47. Statistical tables indicate minimum values for correlation coefficients to be accepted as significant at the .05 and .01 levels.

(2). *Results.* The questions are listed in their rank order and the average percentage response per hour is given for the zero dosage group, 25-75 microgram LSD-25 group and 100-225 microgram LSD-25 group in Tables 3, 4, and 5, respectively. In Table 6 the rank of each question within each of the three dose groups is given in order to enable comparison of the relative importance of the questions in the different groups.

#### Table 6

In the zero dosage group, the four highest ranking questions were (24) "Are your palms moist?", (45) "Do you feel drowsy?", (13) "Does your head ache?", and (47) "Are you anxious?". The lowest ranking items in this group to which no subject responded at any time were the following: (12) "Are your lips drawn back as if you were smiling?", (14) "Are things moving around you?", (23) "Are you cold?", (25) "Are your palms dry?", (36) "Do you have difficulty in focusing your vision?", (37) "Do you see double?", (40) "Do things seem too close?", and (41) "Do things seem too far away?".

In contrast to this, the four questions receiving highest percentage response by the low dose group were: (24) "Are your palms moist?", (16) "Do you feel unsteady?", (29) "Do your hands and feet feel peculiar?", and (42) "Do you tremble inside?", while those receiving the least response were: (10) "Do you have a bitter taste in your mouth?", (18) "Do you pass more urine than usual?", (26) "Are your palms cold?", (34) "Is your hearing more acute than usual?", (40) "Do things seem too close?", (24) "Are your palms dry?", and (37) "Do you see double?".

Among the high dose group, questions (16) "Do you feel unsteady?", (42) "Do you tremble inside?", (24) "Are your palms moist?", and (43) "Do you feel weak?" were most frequently responded to; the following items received the lowest average hourly percentage response: (4) "Is salivation increased?", (18) "Do you pass more urine than usual?", (39) "Does light bother you?", (33) "Is your hearing abnormal?", and (37) "Do you see double?".

TABLE 6  
RANK OF EACH QUESTION FOR EACH OF THREE LSD-25 DOSAGE GROUPS  
(Rank 1 indicates highest percent of positive responses. Succeeding numbers represent decreasing percents.)

Question number	Question	Rank of question		
		N = 13 0 Gamma	N = 21 25-75 Gamma	N = 14 100-225 Gamma
1	Do you feel ill in any way?	10.5	32	15
2	Are you nauseated?	28	24.5	13
3	Have you a feeling of choking?	37.5	32	34
4	Is salivation increased?	8	24.5	44
5	Or decreased?	16	28.5	19
6	Is your appetite increased?	19.5	32	27.5
7	Or decreased?	5.5	26.5	23.5
8	Do you have a "dry" taste in your mouth?	19.5	21	29
9	Do you have a funny taste in your mouth?	10.5	18.5	23.5
10	Is it a bitter taste?	24	43	35
11	Are your lips numb?	37.5	23	20.5
12	Or drawn back as if you were smiling?	43.5	18.5	31.5
13	Does your head ache?	3	11	22
14	Are things moving around you?	43.5	37	27.5
15	Do you feel dizzy?	13	10	5.5
16	Or unsteady?	28	2	1
17	Is there difficulty in breathing?	28	37	40
18	Do you pass more urine than usual?	19.5	43	44
19	Are you aware of your heart beat?	32.5	21	40
20	Is it faster than usual?	19.5	35	36.5
21	Are you sweating?	37.5	21	25.5
22	Are you hot?	13	12	11.5
23	Or cold?	43.5	37	36.5
24	Are your palms moist?	1	1	3.5

TABLE 6 (continued)

Question number	Question	Rank of question		
		N = 13 0 Gamma	N = 21 25-75 Gamma	N = 14 100-225 Gamma
25	Are your palms dry?	43.5	46	40
26	Or cold?	8	43	40
27	Is your skin sensitive?	24	32	25.5
28	Do you have funny feelings on your skin?	32.5	9	7
29	Do your hands and feet feel peculiar?	8	3	5.5
30	Do they feel heavy?	13	13	9
31	Or light?	32.5	14	17.5
32	Is there pressure in your ears?	24	17	15
33	Is your hearing abnormal?	24	28.5	46
34	Is it more acute than usual?	24	43	30
35	Is your eyesight blurred?	32.5	15.5	15
36	Do you have difficulty in focusing your vision?	43.5	26.5	11.5
37	Do you see double?	43.5	47	47
38	Are shapes and colors altered in any way?	37.5	32	31.5
39	Does light bother you?	32.5	39.5	44
40	Do things seem too close?	43.5	43	40
41	Or too far away?	43.5	39.5	33
42	Do you tremble inside?	16	4	2
43	Do you feel weak?	32.5	5.5	3.5
44	Or fatigued?	5.5	5.5	17.5
45	Do you feel drowsy?	2	8	10
46	Do you feel as if in a dream?	16	7	8
47	Are you anxious?	4	15.5	11.5



It can be seen from Table 6 that some of the questions hold similar relatively important positions in all three groups and that several questions are comparatively insignificant for each group. In some cases questions are ranked similarly for the two drug groups, but differ from the zero dosage group. The exact relationship among rank positions of items within three dosage groups was determined by means of a rank difference correlation coefficient.

TABLE 7  
CORRELATION BETWEEN RANK ORDER OF QUESTIONS IN THREE LSD-25 DOSE GROUPS,  
COMPARED TWO AT A TIME  
(N = 47 questions)

LSD-25 Groups Compared	Correlation coefficient*
0 and 25-75 gamma	.46
0 and 100-225 gamma	.36
25-75 and 100-225 gamma	.83

\* Significant correlation coefficient at the .05 level must be .29 and at the .01 level, .37.

Correlation coefficients obtained appear in Table 7. It can be seen that the correlation coefficient between the rank positions of questions in the zero dose group and the 25-75 microgram group was 0.46 and that between the zero group and the high drug group was 0.36. It is important to observe that the ranks of questions in the two drug groups correspond more closely than the ranks of questions in either of the drug groups do with the non-drug groups. In this situation, the correlation coefficient was 0.83. All of these coefficients are statistically significant at the .05 level; all but the 0.36 are significant at the .01 level, and this coefficient practically attains this level of significance.

(3). *Discussion.* The low, but positive, correlations between the zero dose group and each of the drug groups seem to indicate that something other than LSD-25 influenced the response to the questionnaire items. However, the fact that a number of items received little response from subjects in all three groups must not be overlooked. Because of this the ranks of these questions were similarly low and tended to make the correlation positive. The high correlation between item rank of the 25-75 microgram and 100-225 microgram groups shows, however, that the drug does have an effect on physiological and perceptual phenomena which does not appear when a placebo is substituted for the LSD-25. This effect is highly similar in kind at the two doses, but per cent response, intensity of response, and persistency of response may vary. This cannot be determined from a comparison of rank order.

d. *Discriminative questions.*

(1). *Statistical procedure.* In order to determine which of the 47 questions were responded to significantly more often by members of one dose group than by another, a formula for reliability of the difference between percentages was applied to the data.

Formulae are available for both independent and correlated groups. The groups of subjects previously analyzed were only partially correlated since a number of the subjects were tested at only one level. Since the sample was too small to assume that the uncorrelated portions of these groups are drawn from the same population and are therefore alike, correlated groups only were used in this analysis.

Rather than use groups in which dosages of the LSD-25 vary, a more refined technique was used, and only subjects tested at zero dosage, 50 micrograms of the drug, and 100 micrograms were studied. In this way individual variability, which might be counteracted by the use of different doses of LSD-25, would be avoided and a more accurate measure of the significance of the difference could be obtained.

Results obtained from 11 subjects tested at zero and 50 micrograms were compared; seven who were tested under both zero and 100 micrograms of the drug were compared, and comparison of 10 subjects examined at the two drug levels of 50 and 100 micrograms was made.

The formulae applied to the data were:

$$\rho = \frac{\text{Number of subjects giving plus response in first } 3\frac{1}{2} \text{ hours} \times 100}{N} \quad (4)$$

where  $\rho$  equals the percentage of subjects responding positively and  $N$  equals the number of subjects in the group;

$$\sigma_{D\%} = \sqrt{\frac{b + c}{N}} \quad (5)$$

where  $\sigma_{D\%}$  equals the standard deviation of the difference between percentages,  $b$  equals the percentage of subjects responding negatively under the lower dose and positively under the higher, and  $c$  equals the per cent of subjects responding negatively under the higher dose and positively under the lower dose:

$$t = \frac{D_{\%}}{\sigma_{D\%}} \quad (6)$$

where  $D$  equals the difference between the percentages of subjects responding positively under the two doses compared, and  $t$  equals the critical ratio.  $t$  tables indicate minimum values which must be obtained before rejecting the hypothesis that the true difference between percentages responding under different levels is zero.

Equation 4 was used for deriving percentage positive response rather than Equation 2 since it prevents weighting results to make questions appear more reliable. In this way if a subject reports a symptom at each interval it is considered as only one positive response. Responses in the last two hours of questioning were discarded since not all subjects were quizzed at those times.

(2). *Results.* In Table 8 appear the questions discriminating between zero dosage, and 50 microgram dose groups; that is, these are the questions

TABLE 8  
QUESTIONS DISCRIMINATING BETWEEN ZERO AND FIFTY GAMMA LSD-25  
(N = 11)

Questions	Percent responding positively in any hour		t
	Zero LSD-25	50 gamma LSD-25	
	<i>.01 level</i>		
	none	none	none
	<i>.02 level</i>		
16. Do you feel unsteady?	9	91	3.00
46. Do you feel as if in a dream?	18	91	2.83
	<i>.05-.03 levels</i>		
28. Do you have funny feelings on your skin?	9	73	2.65
42. Do you tremble inside?	27	91	2.65
32. Is there pressure in your ears?	9	64	2.46
36. Do you have difficulty in focusing your vision?	0	55	2.46
43. Do you feel weak?	18	73	2.46
31. Do your hands and feet feel light?	9	55	2.24
	<i>.10-.06 levels</i>		
12. Are your lips drawn back as if you were smiling?	0	36	2.00
15. Do you feel dizzy?	36	73	2.00
45. Do you feel drowsy?	46	82	2.00
27. Is your skin sensitive?	9	55	1.89
29. Do your hands and feet feel peculiar?	27	73	1.89
	<i>Approx. .10 level</i>		
4. Is salivation increased?	18	46	1.74
6. Is your appetite increased?	27	55	1.74
21. Are you sweating?	18	46	1.74
23. Are you cold?	9	36	1.74
44. Do you feel fatigued?	55	82	1.74

responded to positively by significantly more people under one dose than under the other. At the .01 level, there are no questions that discriminate. At the .02 level, there are two items: (16) "Do you feel unsteady?", where 9 per cent of the zero dose group and 91 per cent of the 50 microgram group responded positively, giving a  $t$  of 3.00; and (46) "Do you feel as if in a dream?", where 18 per cent of the zero group and 91 per cent of the 50 microgram group responded giving a  $t$  of 2.83. Only two times in one hundred would differences as large as these appear by chance. These were the two most discriminating items between these two dose levels. Between the .03 and .05 levels of significance there were six items which discriminated, and these were: (28) "Do you have funny feelings on your skin?", with a  $t$  of 2.65; (+2) "Do you tremble inside?", with a  $t$  of 2.65; (32) "Is there pressure in your ears?", with a  $t$  of 2.46; (36) "Do you have difficulty focusing your vision?", with a  $t$  of 2.46; (43) "Do you feel weak?", with a  $t$  of 2.46; and (31) "Do your hands and feet feel light?", with a  $t$  of 2.24. Items where  $t$  values have not reached the .05 level are usually not considered significant statistically. When large samples as great as 1000 are used, a  $t$  of 1.96 is necessary for significance at the .05 level; for smaller samples, such as were used in this study, higher values are necessary. However, with the possibility that  $t$  values would not vary greatly if this sample were increased questions which have  $t$  of about 1.96 are enumerated. These are: (12) "Are your lips drawn back as if smiling?", with a  $t$  of 2.00; (15) "Do you feel dizzy?", with a  $t$  of 2.00; (45) "Do you feel drowsy?", with a  $t$  of 2.00; (27) "Is your skin sensitive?", with a  $t$  of 1.89; and (29) "Do your hands and feet feel peculiar?", with a  $t$  of 1.89. Percentages responding to the questions appear in the table. There were five items that were significant at almost the .10 level. These were (4) "Is salivation increased?", (6) "Is your appetite increased?", (21) "Are you sweating?", (23) "Are you cold?", (44) "Are you fatigued?". Thus, symptoms which are related to the ingestion of 50 micrograms of lysergic acid diethylamide, and which occur significantly less frequently under a zero dosage, are, in their order of importance: unsteadiness, dreamlike feelings, funny feelings on the skin, inner trembling, pressure in the ears, difficulty in focusing vision, feeling of weakness, feeling of lightness in hands and feet, lips drawn back as if smiling, dizziness, drowsiness, sensitiveness of skin, and peculiar feelings in hands and feet.

Symptoms which appear to have somewhat of a relationship to the drug but are also present, to a lesser extent, among the subjects of the zero group are: increased salivation, increased appetite, sweating, coldness, and fatigue.

The next table under consideration is Table 9, in which the questions discriminating between zero microgram and 100 microgram of LSD-25 appear.



TABLE 9  
QUESTIONS DISCRIMINATING BETWEEN ZERO AND ONE HUNDRED GAMMA LSD-25  
(N = 7)

Questions	Percent responding positively in any hour		t
	Zero LSD-25	100 gamma LSD-25	
	<i>.02-.01 levels</i>		
	none	none	none
	<i>.05 level</i>		
14. Are things moving around you?	0	86	2.46
16. Do you feel unsteady?	14	100	2.46
28. Do you have funny feelings on your skin?	0	86	2.46
43. Do you feel weak?	14	100	2.46
46. Do you feel as if in a dream?	14	100	2.46
	<i>.10-.06 levels</i>		
1. Do you feel ill in any way?	14	86	2.27
2. Are you nauseated?	0	71	2.27
15. Do you feel dizzy?	29	100	2.27
27. Is your skin sensitive?	0	71	2.27
29. Do your hands and feet feel peculiar?	14	86	2.27
42. Do you tremble inside?	29	100	2.27
21. Are you sweating?	14	71	2.00
31. Do your hands and feet feel light?	0	57	2.00
35. Is your eyesight blurred?	14	71	2.00
36. Do you have difficulty in focusing your vision?	0	57	2.00
41. Do things seem too far away?	0	57	2.00
	<i>Approx. .10 level</i>		
3. Have you a feeling of choking?	0	43	1.73
11. Are your lips numb?	0	43	1.73
17. Is there difficulty in breathing?	0	43	1.73
23. Are you cold?	0	43	1.73
32. Is there pressure in your ears?	0	43	1.73
38. Are shapes and colors altered in any way?	0	43	1.73

Seven subjects were studied. There were no questions which discriminated at better than the .02 level. However, at the .05 level, there were five questions that discriminated. These were: (14) "Are things moving around you?", (16) "Do you feel unsteady?", (28) "Do you have funny feelings on your skin?", (43) "Do you feel weak?", and (46) "Do you feel as if in a dream?". All of these items have critical ratios of 2.46. Among the zero group, no more than 14 per cent, or only one subject, ever responded to these questions, while at least 86 per cent and frequently 100 per cent of those in the drug group responded positively to these questions. Between the .06 and the .10 level, eleven items were significant. The six whose *t* values were 2.27 were: (1) "Do you feel ill in any way?", (2) "Are you nauseated?", (15) "Do

you feel dizzy?", (27) "Is your skin sensitive?", (29) "Do your hands and feet feel peculiar?", and (42) "Do you tremble inside?". A minimum of 71 per cent reported these symptoms when under the drug and a maximum of 29 per cent in the zero group gave positive responses. The remaining five items had a critical ratio of 2.00. They were: (21) "Are you sweating?", (31) "Do your hands and feet feel light?", (35) "Is your eyesight blurred?", (36) "Do you have difficulty in focusing your vision?", (41) "Do things seem too far away?". There were six items which were significant at about the .10 level. All of these had a critical ratio of 1.73, and to these six items, none of the people in the zero group ever responded positively, whereas 43 per cent of the people receiving 100 micrograms of LSD-25 responded positively. These were: (3) "Do you have a feeling of choking?", (11) "Are your lips numb?", (17) "Is there difficulty in breathing?", (23) "Are you cold?", (32) "Is there pressure in your ears?", and (38) "Are shapes and colors altered in any way?".

Some of the symptoms highly significant under 50 micrograms are also significant under 100 micrograms. Included in this category are: unsteadiness, dreamlike feelings, feelings of weakness, funny feelings on the skin. Certain symptoms, namely pressure in ears, difficulty in focusing the vision, feeling of lightness in hands and feet, inner trembling, are most significant for the low dose than the high dose, while a new symptom, things moving about, first becomes significant under 100 micrograms.

Among the questions which are significant at the .06-.10 level for the two groups a number of *new* symptoms occur at the high dose: feeling of illness, nausea, sweating, and blurred eyesight.

A comparison of ten subjects tested with 50 and with 100 micrograms (see Table 10) yields very few significant differences in per cent responses to questions despite the comments just made which compare the two groups in a "non-statistical" way. None of the questions discriminate at the .05 level or better. Between the .06 and .10 levels, two items, with critical ratios of 2.00, are significant; (1) "Do you feel ill in any way?", and (30) "Do your hands and feet feel heavy?". At approximately the .10 level, there are four items; (2) "Are you nauseated?", (9) "Do you have a funny taste in your mouth?", (41) "Do things seem too far away?", and (47) "Are you anxious?". For all of these items, although the differences are not highly significant, more people receiving 100 micrograms gave positive responses than those receiving 50 micrograms of LSD-25.

(3). *Discussion.* Unsteadiness and a dream-like state are the main differentiating symptoms between zero and 50 micrograms. These, it would appear, are due to some interference with higher brain processes and possibly

TABLE 10  
QUESTIONS DISCRIMINATING BETWEEN FIFTY AND ONE HUNDRED GAMMA LSD-25  
(N = 10)

Questions	Percent responding positively in any hour		t
	50 gamma LSD-25	100 gamma LSD-25	
	<i>.05-.01 levels</i>		
	none	none	none
	<i>.10-.06 levels</i>		
1. Do you feel ill in any way?	30	70	2.00
30. Do your hands and feet feel heavy?	30	70	2.00
	<i>Approx. .10 level</i>		
2. Are you nauseated?	40	70	1.76
9. Do you have a funny taste in your mouth?	20	50	1.76
41. Do things seem too far away?	20	50	1.76
47. Are you anxious?	50	80	1.76

involve the ascending reticular pathways, described by Magoun (1) and his workers. Symptoms which discriminate between the .03 and .05 levels (paraesthesias, inner trembling, pressure in the ears, difficulty in focusing, weakness, and lightness of the limbs) are predominantly perceptual changes, and are probably again central in origin and not peripheral. Between zero and 100 micrograms the most discriminating items are those describing things moving around, unsteadiness, paraesthesias, weakness, and dream-like state. They are quite similar in nature to the items differentiating 50 micrograms from zero, with the exception of "things moving around you." Apparently it takes a higher dosage of LSD-25 to elicit vertiginous symptoms. Higher doses of the drug are also required before feelings of illness, nausea, funny taste in mouth, and anxiety are significant. Heaviness of the limbs and objects seeming too far away also occur more frequently under 100 micrograms.

The number of subjects analyzed has considerable influence upon the significance of results obtained. In view of results obtained with small samples, it is likely that more symptoms would be of greater statistical significance if the *N* were increased. These data, however, indicate the *most reliable* symptoms since they showed differences even with small samples.

While there are not too many items discriminating between 50 and 100 micrograms of LSD-25, and these are not at high levels of significance, it is suggested that the subjective severity of symptoms is greater under the high dose than under the 50 microgram dose. The original response to the questionnaire was made on a one-plus to a five-plus basis and analysis of these responses to questions answered positively by both groups might indicate

differences in severity. The number of subjects used here is too small to warrant this break-down.

Symptoms under 100 micrograms might be experienced more rapidly after ingestion of the drug and might be more persistent than those under 50. The hour-by-hour analysis of responses gave some indication of this. We shall later show that the mean number of symptoms reported each hour tends to support this theory.

Among the items which did not discriminate between dose groups were items to which either a large percentage response or a small percentage response was made by both groups being compared. The former symptoms are probably independent of the pharmacological action of the drug and could be a result of stress and anxiety, environmental conditions, suggestibility of subjects, or some other unknown variable. Only the symptoms which are reported significantly more frequently upon the actual administration of the drug may be ascribed to the action of the drug itself, either directly or indirectly.

## 2. Number Analysis

### a. Number of symptoms reported: Comparison of dosage groups.

(1). *Statistical procedure.* The mean number of symptoms reported during each of the six question periods by subjects in each of the three dosage groups was computed by the following formula:

$$M_{H_i} = \frac{\sum_{i=1}^N + H}{N} \quad (7)$$

where  $M_{H_i}$  equals the mean number of symptoms in a given hour,  $\sum_{i=1}^N + H$  equals the total number of questions responded to positively in a given hour by all subjects, and  $N$  equals the number of subjects. Mean number of symptoms reported per hour during the  $5\frac{1}{2}$  hour test period was determined with the formula:

$$M_{PH} = \frac{\sum_{i=1}^6 M_{H_i}}{6} \quad (8)$$

where  $M_{PH}$  equals the mean number of symptoms per hour,  $\sum_{i=1}^6$  equals the sum of Items 1 through 6, and  $M_{H_i}$  equals the mean number of symptoms reported each hour, as calculated with Equation 7.



In order to determine whether differences between the number of symptoms reported by the various dose groups were significantly different from zero and would not have occurred by chance more than five times in a hundred, it was necessary first to compute the mean number of symptoms reported by the same subjects tested at two dose levels. In the comparison between zero and 25-75 micrograms LSD-25, 11 subjects were available for analysis; there were 11 tested at 25-75 micrograms and 100-225 micrograms who were compared, and five whose results under zero and under 100-225 micrograms of the drug were studied. If a subject had been tested at two or more levels within a group, only the data from the lowest dose were used; if a subject had been tested more than once at a given dose, only data from the first testing were used. Equations 9, 10, 11, and 12 were used.

$$M_D = \frac{\sum_{i=1}^N d_i}{N} \quad (9)$$

where  $M_D$  equals the mean difference between the number of symptoms reported by groups tested at two levels,  $d_i$  equals the difference between the number of symptoms reported by a given subject tested at two levels, and

$\sum_{i=1}^N$  equals the sum of 1 to  $N$  individuals.

$$SD_D = \sqrt{\frac{\sum x^2}{(N-1)}} \quad (10)$$

where  $SD_D$  equals the standard deviation of the difference,  $\sum x^2$  equals the sum of the squared deviations of each subject's difference from the mean difference, and  $N$  equals the number of subjects.

$$SE_{M_D} = \frac{SD_D}{\sqrt{N}} \quad (11)$$

where  $SE_{M_D}$  equals the standard error of the mean difference,  $SD_D$  equals the standard deviation of the difference as obtained with Formula 10.

$$t = \frac{M_D}{SE_{M_D}} \quad (12)$$

where  $t$  equals the critical ratio,  $M_D$  equals the mean difference calculated with formula (9), and  $SE_{M_D}$  equals the standard error of the mean difference, calculated with formula (11).

Statistical *t* tables indicate minimum *t* values which must be obtained to warrant rejection of the hypothesis that the real difference between groups is zero.

(2). *Results.* From Table 11 it can be seen that for a group of 11 subjects tested with zero LSD-25 the maximum average number of positive responses occurred  $\frac{1}{2}$  hour after the placebo, when 4.9 responses were given. The number of reported symptoms gradually declines thereafter, except for a slight rise at  $3\frac{1}{2}$  hours. An average of 3.1 questions per hour received positive responses. It should be noted that one of the 11 subjects was not asked the questions at  $4\frac{1}{2}$  and  $5\frac{1}{2}$  hours after the drug.

As seen in Table 12, responses made by a group of five subjects given placebos are somewhat similar to those of the groups summarized in the previous table. The maximum response was again given in the first  $\frac{1}{2}$  hour

TABLE 11  
COMPARISON OF AVERAGE NUMBER OF SYMPTOMS UNDER ZERO LSD-25 AND  
25-75 GAMMA LSD-25  
(Groups are compared at each of six question periods and for total six periods)

Hours after drug	Average number symptoms Gamma LSD-25		<i>t</i> *	N
	Zero	25-75		
$\frac{1}{2}$	4.9	4.3	.44	11
$1\frac{1}{2}$	4.1	13.2	4.11	11
$2\frac{1}{2}$	2.3	13.5	7.03	11
$3\frac{1}{2}$	2.9	12.0	4.46	11
$4\frac{1}{2}$	1.8	9.5	4.05	10
$5\frac{1}{2}$	1.8	10.0	3.34	10
Average $5\frac{1}{2}$	3.1	10.6	5.71	11

\* See Table 14 for values of "*t*" at the different levels of significance.

TABLE 12  
COMPARISON OF AVERAGE NUMBER OF SYMPTOMS UNDER ZERO LSD-25 AND 100-225  
GAMMA LSD-25  
(Groups are compared at each of six question periods and for total six periods)  
N = 5

Hours after drug	Avg. No. Symptoms Gamma LSD-25		<i>t</i> *
	Zero	100-225	
$\frac{1}{2}$	3.8	9.4	1.99
$1\frac{1}{2}$	1.8	19.0	6.25
$2\frac{1}{2}$	3.0	18.2	4.75
$3\frac{1}{2}$	2.6	16.2	3.22
$4\frac{1}{2}$	1.8	14.2	2.29
$5\frac{1}{2}$	1.8	12.4	4.17
Average $5\frac{1}{2}$	2.5	14.9	6.10

\* See Table 14 for values of "*t*" at the different levels of significance.

when an average of 3.8 questions were answered positively. After the decrease at 1½ hours there was an increase, after which the number of symptoms declined. The average number of symptoms throughout the day was 2.5, which is .6 symptoms less than for the group of 11 subjects.

The average number of symptoms reported per hour by the 25-75 microgram group was 10.6 for an "N" of 11 (see Table 11) and 8.9 (see Table 13) for another group of 11 subjects some of whom were the same. The peak hour was at 2½ hours for one group and 1½ hours for the other. In each group, however, the difference between symptoms reported at 1½ and 2½ hours is negligible and the time of greatest reaction to the drug may be thought of as extending from 1½ to 2½ hours after the drug. A decrease in positive responses followed for both groups, with a slight increase at 5½ hours.

TABLE 13  
COMPARISON OF AVERAGE NUMBER OF SYMPTOMS UNDER 25-75 GAMMA LSD-25 AND  
100-225 GAMMA LSD-25  
(Groups are compared at each of six question periods and for total six periods)

Hours after drug	Avg. No. Symptoms		t*	N
	25-75 Gamma	100-225 Gamma		
½	5.4	12.0	3.13	11
1½	11.9	16.7	3.77	11
2½	11.8	15.1	1.64	11
3½	9.8	14.6	2.16	10
4½	6.9	13.3	2.95	10
5½	10.8	13.8	1.07	4
Average 5½	8.9	13.6	4.48	11

\* See Table 14 for values of "t" at the different levels of significance.

Under 100-225 micrograms five subjects gave an average 14.9 positive responses per hour, while a group of 11 reported an average of 13.6 symptoms each hour. In both groups the peak hour occurred 1½ hours after the drug; following this peak there was a very slight decline. By the last question hour as many as 12.4 and 13.8 positive responses were still being made. This indicates that the drug effects may persist for a more extended period with high dosage.

Despite the variability between each pair of the three dose groups, the number of positive responses is correlated with the amount of drug; increasing amounts of drug and greater number of reported symptoms, out of the 47 suggested ones, tended to occur together. A statistical analysis of the differences between the number of positive responses made each hour by each group indicate whether the differences would have occurred by chance or whether they are true differences. The t values, or critical ratios, reported in Tables 11-13 may be interpreted with the aid of Table 14. The minimum

values of  $t$  which must be obtained before a difference for a given sample size can be considered significant are given in this table.

TABLE 14  
 "t" VALUES AT FOUR LEVELS OF SIGNIFICANCE  
 (Values are given for various "N's" used in comparisons made in Tables 11, 12, and 13)

N	Degrees of freedom	.10	levels of significance		
			.05	.02	.01
			"t" values		
4	3	2.35	3.18	4.54	5.84
5	4	2.13	2.78	3.75	4.60
7	6	1.94	2.45	3.14	3.71
10	9	1.83	2.26	2.82	3.25
11	10	1.81	2.23	2.76	3.17
$\infty$	$\infty$	1.65	1.96	2.33	2.58

In comparing 0 with 25-75 micrograms the differences between symptoms reported at each period but the first are significant and would have occurred by chance less than one time in a hundred (see Table 11). While a minimum  $t$  of 3.17 was necessary for .01 level significance, all values were above this. The most reliable difference appeared at 2½ hours, the peak hour for the 25-75 group.

When the high drug group is compared to the zero dose group, the differences at the ½ and 4½ hour intervals were not significant (at the .05 level). However, all the other hours showed differences significant at the .05 level or better. Here too, the most significant differences were seen at the peak hour of the drug group which was 1½ hours. These results are seen in Table 12.

Less striking and less reliable differences are observed in analyzing the 25-75 and 100-225 microgram groups. Although the average number of symptoms per hour show a difference reliable at better than the .01 level, the only significant hours (as shown in Table 14) are the first ½ hour (significant at the .05 level), 1½ hours (significant at better than the .01 level) and 4½ hours (significant at better than the .02 level, but not at the .01 level). The required "t" values are given in Table 14.

The foregoing results indicate that LSD-25 does have an effect upon the physiological and perceptual status of the individual, with greater amounts of drug producing greater effect, as determined by the number of positive responses to the 47 items on the questionnaire. About three symptoms on the average, are reported under a placebo, about 10 under 25-75 micrograms of the drug, and about 14 under 100-225 micrograms (Tables 11, 12, and 13).

(3). *Discussion.* Although the standard deviation of scores about the



mean is not reported in the tables, there is a certain scatter about the mean; consequently when a different group of subjects is tested it is not likely that the identical means will be obtained for each group examined under comparable conditions. This was observed in results obtained. However, despite the differences between means for the same dose group, the differences are small when compared to the differences between dose groups.

The statistics indicate that the high dose of drug takes effect as early as  $\frac{1}{2}$  hour after ingestion, while the low dose gives responses similar to the zero dose group at that hour. In addition to its more rigid effectiveness the high dose produces more persistent symptoms. The average number of symptoms reported by the 100-225 group at each hour is greater than the maximum number of the low dose group, as shown in Table 13. At  $2\frac{1}{2}$  hours, the difference between these two groups is not significant. In the low group, the peak seems to occur between  $1\frac{1}{2}$  and  $2\frac{1}{2}$  hours, with little difference between number of symptoms at these two periods. The high group peak is definitely at  $1\frac{1}{2}$  hours, and declines thereafter. Thus the difference at  $2\frac{1}{2}$  hours becomes insignificant. The last hour ( $5\frac{1}{2}$ ) also shows no reliable difference. However, there were only four subjects questioned at this hour, and thus the means are less reliable than for the other hours, and this comparison deserves less consideration.

The increase in number of symptoms reported by the 25-75 microgram group in the last hour, as shown in the group compared with the zero dose group (Table 11) appears to be too small to be significant. A larger increase is observed at this hour when the 25-75 and 100-225 microgram groups are compared. However, only four subjects were tested at this level, in contrast to 10 or 11 subjects examined during the previous hours. The results for the last hour are not comparable to the rest of the results and the rise in mean number of symptoms may be disregarded.

Despite the observation earlier in this report that only a relatively small number of the 47 items on the questionnaire were significant items, it is interesting that comparison among groups, of number of symptoms reported shows definite differences. With increasing dosage there is also an increased tendency to respond positively. It is suggested that if the responses made to only the significant questions were analyzed in the same way, far greater differences having far greater significance would be found.

*b. Relationship between subjects responding at different doses.*

(1). *Statistical procedure.* To investigate the significance of the individual subject the rank difference correlation coefficient is computed with Formula (3) given previously:

$$\rho = 1 - \frac{6\sum D^2}{N(N^2 - 1)} \quad (3)$$

where  $\rho$  equals the rank difference correlation coefficient,  $D$  equals the difference between the rank of an individual in one dose group and his rank in another group; rank 1 is delegated the subject reporting the fewest mean number of symptoms per hour, etc.; and where  $\sum D^2$  equals the sum of the squared differences in rank for all subjects in the groups compared, and  $N$  equals the number of subjects.

Once again the three dose groups were compared two at a time, with only correlated groups used. The number of subjects in each pair of groups was as follows: (a) zero and 25-75 micrograms LSD-25 (11); (b) zero and 100-225 micrograms LSD-25 (5); (c) 25-75 micrograms and 100-225 micrograms LSD-25 (11).

If subjects were tested at two levels in a given dose group, data from only the lowest level were used; if a subject was tested more than once at a given level, data from only the first testing were used.

(2). *Results.* The rank difference correlation coefficients of the average number of symptoms reported per hour by subjects tested at three different dose levels, compared two at a time are given in Table 15. Between the zero

TABLE 15  
CORRELATION BETWEEN AVERAGE NUMBER OF SYMPTOMS REPORTED PER HOUR BY  
SUBJECTS AT 3 LSD-25 DOSE LEVELS

LSD-25 Dose Groups compared	N	Correlation coefficient	Significant Corr. Coef.	
			.05 level	.01 level
zero and 25-75 gamma	11	.66	.60	.74
zero and 100-225 gamma	5	.60	.88	.96
25-75 and 100-225 gamma	11	.90	.60	.74

and 25-75 microgram groups there were 11 subjects and the correlation between their ranks in the two groups (according to the number of symptoms they reported) was .66. This is significant at somewhere between the .01 and .05 level. Between zero and 100-225 micrograms, where there were only five subjects, the correlation was .60. For this number of subjects, the correlation is not significant at the .05 level. With a greater number of subjects, significance would probably be reached. Between the 25-75 and 100-225 microgram level, there was a correlation of .90 for 11 subjects; this is significant at better than the .01 level.

(3). *Discussion.* The results obtained in this analysis are highly important in indicating the significance of the individual subject. The extremely

high, almost perfect, correlation between results of the two groups that took LSD-25 indicates that the person reporting relatively little effect under an intermediate dose will also report relatively little effect under a high dose. Similarly, the person in the group who responds positively to more questions than other subjects tested at a low dose will respond similarly at a high dose.

Previously discussed data showed that a reliably greater number of symptoms (see Table 13) was reported under the high dose than under the low dose. The results reported in Table 15 show that despite the fact that subjects gave more positive responses under 100-225 micrograms, the response made seems to be a function of another factor in addition to LSD-25.

The high correlation between an individual's rank position at zero dose and each of the two drug groups has even more interesting implications. While the correlations are not perfect, they are sufficiently high to permit a fair amount of accuracy in predicting whether an individual will show a small or large reaction under the drug, on the basis of his reaction under zero dosage. There are apparently unknown factors in a given individual which make him more or less suggestible at each level of dosage.

### c. Consistency of response under same dose.

(1). *Statistical procedure.* The Pearson product-moment correlation coefficient was calculated to show the relationship between results obtained under two testings of a given subject under the same dose. The number of symptoms reported at each of the six question intervals and the average number of symptoms per hour under the two testings were compared by Equation (13).

$$r = \frac{N\sum XY - \sum X \times \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} \quad (13)$$

where  $r$  equals the coefficient of correlation,  $N$  equals the number of subjects,  $X$  equals the number of symptoms reported during the first testing, and  $Y$  equals the number of symptoms reported during the second testing. Nine subjects were studied in this manner. While each subject was tested twice at a given dose, the doses ranged from 25 micrograms to 200 micrograms of the drug. One subject was tested twice at each of three dosages; his responses at the first testing under each dose were average, and those of the second testings were averaged.

(2). *Results.* The data tabulated in Table 16 gives some measure of the reliability of responses made to the questionnaire under the influence of lysergic acid diethylamide. The closer the correlation coefficient is to 1.00,

TABLE 16  
CORRELATION BETWEEN NUMBER OF SYMPTOMS REPORTED AT TWO SEPARATE TESTINGS  
UNDER SAME LSD-25 DOSAGE  
(N = 9)

Hours after LSD	Correlation coefficient*
$\frac{1}{2}$	.34
$1\frac{1}{2}$	.42
$2\frac{1}{2}$	.45
$3\frac{1}{2}$	.63
$4\frac{1}{2}$	.58
$5\frac{1}{2}$	.72
Total $5\frac{1}{2}$	.77

\* Correlation coefficients must equal .67 to be significant at the .05 level, and .80 to be significant at the .01 level.

the more likely it is that second testings under the same dose will elicit the same number of responses. The following coefficients were obtained:

At the first  $\frac{1}{2}$  hour, there was a correlation coefficient of 0.34; at the  $1\frac{1}{2}$  hour interval it was 0.42; at  $2\frac{1}{2}$  hours, it was 0.45; at  $3\frac{1}{2}$  hours, it was 0.63; at  $4\frac{1}{2}$  hours, it was 0.58; and at  $5\frac{1}{2}$  hours, 0.72. Thus, the correlation coefficient tends to increase with increasing time after administration of LSD-25. Only the coefficient at  $5\frac{1}{2}$  hours after the drug is significant at the 0.5 level.

The correlation coefficient between the total number of symptoms reported under the two testings was .77. This is significant at the .05 level and almost reaches significance at the .01 level. Thus, on the whole, it can be said that the number of responses to a specified dose is fairly constant for each individual.

(3). *Discussion.* It seems unusual that a correlation coefficient as high as .72 would be obtained for the last testing hour, when the previous hours showed lower correlations. However, this is probably a function of the smaller number of symptoms reported in the last hour.

The correlations obtained at each of the other hourly intervals might be more reliable and higher if a greater number of subjects were used. The fact that the correlation for the total number of symptoms during the day is significant (and higher) tends to support this hypothesis, since this figure was based on a larger number of observations.

The fact that only some of the questions in the questionnaire are significant questions in terms of LSD-25 effects, might contribute to the lack of reliable and high correlation coefficients. Recalculation of these coefficients in terms of the number of discriminating questions reported at the two testings might produce results indicating greater consistency in reporting "reliable" symptoms.



d. Relationship between body weight and symptoms reported.

(1). *Statistical procedure.* The subject's reported body weight and the number of questions to which he responded positively at each of three LSD-25 doses (0 micrograms, 50 micrograms, and 100 micrograms) were compared by Equation 13a for periods of  $\frac{1}{2}$  hour and  $1\frac{1}{2}$  hours after the drug to determine the relationship between the two.

$$r = \frac{N\sum XY - \sum X \times \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2]}} \quad (13a)$$

where  $r$  equals the coefficient of correlation,  $N$  equals the number of subjects,  $X$  equals the figure representing the subject's body weight, and  $Y$  equals the number of symptoms reported by subject of body weight  $X$ . The standard deviation of  $r$  was then computed by the formula:

$$\sigma_r = \frac{1 - r^2}{\sqrt{N - 1}} \quad (14)$$

where  $\sigma_r$  equals the standard deviation of correlation coefficient,  $r^2$  equals the squared correlation coefficient, and  $N$  equals the number of subjects.

Appropriate statistical tables indicate minimum coefficients which must be obtained under different sample sizes before the coefficients are considered reliably different from zero.

(2). *Results.* In Table 17, the correlations between the number of symptoms reported at a given dosage and the subject's body weight in pounds are given. At zero dosage, for the first  $\frac{1}{2}$  hour, the correlation is 0.20, with a standard deviation of 0.28; this correlation is not significantly different from zero, for an  $N$  of 13, as indicated in the table. At this same dosage, but  $1\frac{1}{2}$

TABLE 17  
CORRELATION BETWEEN NUMBER OF SYMPTOMS REPORTED AND SUBJECT'S BODY WEIGHT  
(Zero, 50, and 100 gamma are analyzed for  $\frac{1}{2}$  and  $1\frac{1}{2}$  hours after LSD-25)

Amt. LSD-25	Hours after LSD-25	$r$	$\sigma_r$	$N$	Significant correlation coefficients	
					.05 level	.01 level
zero	$\frac{1}{2}$	.20	.28	13	.55	.68
zero	$1\frac{1}{2}$	.49	.22	13	.55	.68
50 gamma	$\frac{1}{2}$	.02	.26	16	.50	.62
50 gamma	$1\frac{1}{2}$	.18	.25	16	.50	.62
100 gamma	$\frac{1}{2}$	.39	.26	12	.58	.71
100 gamma	$1\frac{1}{2}$	.17	.29	12	.58	.71
				1000	.06	.08

hours after the placebo, the correlation is 0.49, the standard deviation 0.22; this almost reaches significance at the .05 level. At 50 micrograms, the correlation between body weight and the number of symptoms reported at the first  $\frac{1}{2}$  hour is 0.02, ( $\pm .26$ ) which is not significantly different from zero. At  $1\frac{1}{2}$  hours, the correlation is 0.18 with a standard deviation of 0.25, so this is not significant. Under 100 micrograms of LSD-25 the correlation is 0.39, ( $\pm .26$ ) for the first  $\frac{1}{2}$  hour;  $1\frac{1}{2}$  hours after the drug the correlation is 0.17, ( $\pm .93$ ). None of these correlations is significant even at the .05 level. If a larger number of subjects were used it is possible that some of these correlations, if they remained the same, would become significant. For an  $N$  of 1000, it would only be necessary to have a correlation of .08 to reach the .01 level, or a correlation of .06 to reach the .05 level of significance. All the obtained correlations, but one, are higher than these values. None of the correlations were significantly different from zero for the number of subjects evaluated. It can be said that there is no relationship between the subject's body weight and the number of symptoms he reports on the questionnaire.

(3). *Discussion.* The literature on experiments investigating lysergic acid diethylamide symptomatology frequently reports use of dosages in proportion to body weight; generally one microgram of drug per kilogram of body weight has been given "normals." This practice is based on the theory of pharmacologic drug action that a larger and heavier person would have a smaller concentration of the drug than a smaller person receiving the same amount of drug if the substance were evenly dispersed through the body.

The correlation coefficients obtained fail to justify this procedure. Although increased sample size might yield different results, coefficients would have to be negative to support the theory of relatively less drug effect upon heavier persons. It is highly unlikely that results from larger samples would become greater in absolute value, and be opposite in direction. The obtained correlation coefficients are all positive and thus, if a larger  $N$  made them significant, they would indicate that heavier subjects report a greater rather than a smaller number of symptoms.

It is of interest to note that the highest and most significant correlation was obtained when only a placebo was given the subjects, suggesting that heavier persons are more suggestible. Whether weight is directly responsible or not cannot be determined.

#### D. SUMMARY AND CONCLUSIONS

On the basis of tests performed on 26 non-psychotic, intelligent adults, who were given from one to three doses of lysergic acid diethylamide (zero, 25-75 micrograms, and 100-225 micrograms), and to whom a questionnaire, in-

vestigating changes in physiological and perceptual phenomena, was given at hourly intervals, the following conclusions are drawn:

1. Symptoms most significantly related to the ingestion of 50 micrograms of the drug are (in order of decreasing significance): unsteadiness, dream-like feeling, paraesthesias, inner trembling, pressure in ears, difficulty in focusing vision, weakness, lightness of limbs, lips drawn back as if smiling, dizziness, drowsiness, sensitivity of skin, and peculiar feeling of limbs. Less significant, but probably related, are: increased salivation, increased appetite, sweating, cold, and fatigue.

2. Symptoms most significantly related to the ingestion of 100 micrograms of the drug are (in order of decreasing significance): things moving about subjects, unsteadiness, paraesthesias, weakness, dream-like feeling, illness, nausea, dizziness, sensitivity of skin, peculiar feeling of limbs, inner trembling, sweating, lightness of limbs, blurred eyesight, difficulty in focusing vision, and objects seeming too far away. The less significant but probably related symptoms are: feeling of choking, numbness of lips, difficulty in breathing, cold, pressure in ears, and alteration of shapes and colors.

3. There may be differences in subjective severity and quality of the symptoms which are reported under both drug levels. However, several symptoms discriminate between 50 and 100 micrograms in that they received a greater percentage of response under the high dose. These are: feeling of illness, heaviness of limbs, nausea, funny taste in mouth, objects seeming too far away, and anxiety.

4. There is a significant correlation of .83 between the relative position of symptoms, according to frequency of positive response, for the two drug groups.

5. The average number of symptoms, out of 47 suggested ones, under zero dosage is about 3; under 25-75 micrograms it is approximately 10; and under 100-225 micrograms it is about 14. The differences among the three groups are all statistically reliable at better than the .01 level of significance.

6. The peak effect under zero dosage occurs in the first  $\frac{1}{2}$  hour, and that of the low dose between  $1\frac{1}{2}$  and  $2\frac{1}{2}$  hours after the drug. The peak for the high dose occurs  $1\frac{1}{2}$  hours after the drug, and the effect is longer lasting than for either of the other groups. These statements are based on group results and individual variability is not considered here.

7. The number of symptoms a subject reports under the low dose correlates .90 with the number he reports under the high dose. Although the average number of symptoms increases, he maintains his relative position within the group. There is also a relationship as high as .66 and .60 between subjects' relative position under zero and under low and high doses of LSD-25, respectively. This indicates a fairly high degree of predictability of the number of

responses to the drug on the basis of the number of responses under the placebo.

8. The reliability of responses to the questionnaire has been found high. Test-retest correlation as high as .77 was obtained in comparing the total number of symptoms reported at two separate testings under the same dosage.

9. The number of symptoms reported and the subject's body weight have been shown either to be unrelated or not related in the expected direction.

### REFERENCES

1. BUSCH, A. K., & JOHNSON, W. C. L.S.D. 25 as an aid in psychotherapy. *Dis. Nerv. Syst.*, 1950, **11**, 2-4.
2. CONDRAU, G. Klinische Erfahrungen an Geisteskranken mit Lysergsäure-Diethylamid. *Acta Psychiat. et Neurol.*, 1949, **24**, 9-32.
3. DESHON, H. J., RINKEL, M., & SOLOMON, H. Mental changes experimentally produced by L.S.D. *Psychiat. Quart.*, 1952, **26**, 33-53.
4. HOCH, H., CATTELL, J. P., & PENNES, H. H. Effects of drugs. *Am. J. Psychiat.*, 1952, **108**, 585-589.
5. HOCH, P. H., CATTELL, J. P., & PENNES, H. H. Effects of mescaline and lysergic acid (d-LSD-25). *Am. J. Psychiat.*, 1952, **108**, 579-584.
6. MAGOUN, H. W. An ascending reticular activating system in the brain stem. *A. M. A. Arch. Neurol. & Psychiat.*, 1952, **67**, 145-154.
7. MATEFI, L. Mezcalin und Lysergsäurediethylamid-Rausch. *Confinia Neurol.*, 1952, **12**, 146-177.
8. MAYER-GROSS, W. Experimental psychoses and other mental abnormalities produced by drugs. *Brit. M. J.*, 1951, **2**, 317-321.
9. RINKEL, M., DESHON, H. J., HYDE, R. W., & SOLOMON, H. C. Experimental schizophrenia-like symptoms. *Am. J. Psychiat.*, 1952, **108**, 572-578.
10. Sandoz Pharmacology Labs., Ltd. D-lysergic acid diethylamide (LSD 25), 1952, 1-18.
11. SAVAGE, C. Lysergic acid diethylamide (LSD-25): A clinical-psychological study. *Am. J. Psychiat.*, 1952, **108**, 896-900.
12. STOLL, W. A. D-lysergic acid diethylamide (LSD-25), a phantasticum of the ergot group. *Schweiz. Arch. f. Neurol. u. Psychiat.*, 1947, **60**, 1-45.

*The Mount Sinai Hospital\**

*Fifth Avenue & 100 Street*

*New York City 29*

---

\*Dr. Kornetsky was assigned to the project by the National Institute of Mental Health, Bethesda, Maryland. He was stationed at the Biological Laboratory, Cold Spring Harbor, Long Island, New York.

We are indebted to the Geschickter Fund for Medical Research, Inc., for its generous support of this project. Thanks are due Dr. M. W. Ferguson who aided in this investigation in its earlier phases.



## STUDIES IN PHYSIOGNOMIC PERCEPTION: II. EFFECT OF DIRECTIONAL DYNAMICS OF PICTURED OBJECTS AND OF WORDS ON THE POSITION OF THE APPARENT HORIZON\*<sup>1</sup>

*Department of Psychology, Clark University*

STANLEY E. KADEN, SEYMOUR WAPNER, AND HEINZ WERNER

### A. THE PROBLEM

In a previous article (9) it was demonstrated that visual dynamics, inherent in visual forms such as the picture of a bird or airplane in flight, exists as a behaviorally measurable event. Directional dynamics of a figure to the left or to the right has been shown to have a significant and differential effect on spatial localization of a figure. In particular, a figure is seen displaced relatively to the left or right depending on whether the directional dynamics is to the left or to the right. This effect was measured by determining the physical location of the so-called apparent median plane (straight-ahead) under variation of direction of dynamics. The apparent median plane shifts opposite the directional dynamics of the figure. If, for instance, a figure with directional dynamics to the left is placed in the objective median plane, i.e., physically straight-ahead, it is seen relatively displaced to the left, and if the subject is required to move it so that it appears to him as straight-ahead, he will have to move it to the right; in other words, the apparent median plane or straight-ahead has shifted to the right because of the dynamic properties inherent in the figure.

The following two experiments have been designed to test the generality of these findings concerning the behavioral effect of directional dynamics: the first experiment discussed below deals with the efficacy of directional dynamics of a configuration on localization in the direction of up and down rather than left and right; the second experiment aims at evaluating whether directional dynamics connoted by verbal symbols (words) have an effect on localization in space analogous to that already demonstrated with pictorial material.

\* Received in the Editorial Office on August 2, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was supported by research grant MH-348 C from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

Both experiments dealt with dynamics in the "up-down" direction, and measured spatial displacements with respect to the horizon (eye line). One has to distinguish between two concepts of the horizon, namely, the objective horizon which is physically defined and the apparent horizon which is perceptually defined. The objective horizon is defined as a horizontal plane bisecting both eyes (under normal forward gaze) into symmetrical upper and lower hemispheres. The apparent horizon, on the other hand, refers to that physical position of a horizontal line in space experienced as at the horizon or eye level.

The two experiments had the following apparatus and procedure in common. The tests were performed in a dark room. *S* sat erect throughout all trials in a chair with side supports; his head was held in position by an adjustable head rest and chin rest; in this way, position of head and trunk were kept constant throughout the trials. In all three experiments luminous test figures were used. The figures were placed on the front face of a light box; this box in turn was mounted on a rack-and-gear so that it was possible to move the light box in small steps up and down along a vertical track in the fronto-parallel plane at a distance of 200 cm. from *S*. The objective horizon was ascertained separately for each *S*; this was achieved through measuring the height of *S*'s pupils with regard to the floor.

The position of the apparent horizon was measured in terms of deviation from objective horizon. Arbitrarily, positions above the objective horizon are designated by plus values; positions below the objective horizon are designated by minus values.

The *S*'s task, in general, was to fixate part of the test figure and to instruct *E* how to move the figure so that the fixated part appeared at the horizon. In each trial the fixated part was initially placed at the objective horizon of the particular *S*.

## B. EXPERIMENT 1: EFFECT OF DIRECTIONAL DYNAMICS OF PICTURED OBJECTS ON THE PHYSICAL POSITION OF THE APPARENT HORIZON

### 1. *Material and Hypothesis*

This experiment, as mentioned above, represents an extension of the study concerned with left-right directional dynamics of a figure to the efficacy of dynamics in the "up-down" direction. For this purpose, pictures of hands pointing upwards, and of hands pointing downwards were used. Such figures are expected to be perceived as displaced in the direction of the dynamics; accordingly, the apparent horizon is expected to shift opposite the direction of the dynamics.

The general hypothesis of this experiment can therefore be formulated as follows: With pictures of hands pointing upwards the apparent horizon will shift relatively more downwards than with hands pointing downwards.

## 2. Procedure

Two luminous silhouettes served as test patterns. As can be seen in Figure 1, they consisted of two hands, one following the other. The *S* was requested

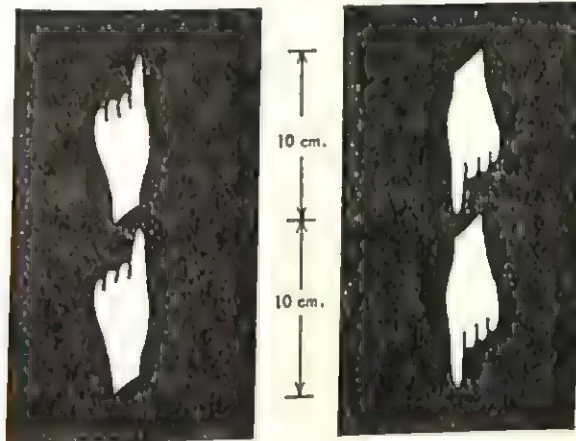


FIGURE 1  
EXPERIMENT 1: TEST PATTERNS

to fixate the finger tip in the middle of the test pattern, and to instruct *E* to move the test pattern until it appeared at eye level.<sup>2</sup>

Each *S* was tested once with each test pattern. Two sequences were employed: half of the *Ss* were presented with hands pointing upward first; half with hands pointing downward first. Twenty *Ss* were tested; 5 men and 5 women in each of the two sequences.

## 3. Results

A summary of the results is presented in Table 1. With hands pointing upwards the mean position of the apparent horizon is relatively downwards,  $-4.28$  cm., as compared with hands pointing downwards,  $-.25$  cm.<sup>3</sup> This

<sup>2</sup>In previous studies (2, 6, 8) it has been demonstrated that extension of a test pattern asymmetrically to one side of fixation significantly affects the physical position of the apparent median plane, and analogous effects have been obtained for the apparent horizon; viz., a luminous rectangle with its base in the objective horizon tends to shift the apparent horizon upwards and vice versa. In order to neutralize such effects of asymmetrical placement the two-hand configuration was chosen and *S* requested to fixate the middle part so that the figure extended equidistantly above and below fixation.

<sup>3</sup>As one will note, for both test conditions negative values were obtained; in this

TABLE 1  
EXPERIMENT 1: EFFECT OF DIRECTIONAL DYNAMICS OF PICTURED OBJECTS ON THE  
PHYSICAL POSITION OF THE APPARENT HORIZON

Source of variation	df	Mean square	F	P
Subjects	19	94.89	—	—
Sequence	9	.03	<1.00	>.05
Sex	1	.76	<1.00	>.05
Subjects within sex and sequence	9	198.10	14.93	<.01
Conditions	1	162.81	12.28	<.01
Order	1	4.56	<1.00	>.05
Error	18	13.26		
Total	39			

<i>Hands Pointing</i>	<i>Mean (cm.)*</i>
Up	-4.28
Down	- .25

\* Positions above the objective horizon are designated by plus values; positions below the objective horizon are designated by minus values.

difference is significant at the .01 level. Thus, the results are in keeping with the hypothesis that there is a relative shift of the physical position of the apparent horizon opposite the direction of the dynamics of the figure.

#### C. EXPERIMENT 2: EFFECT OF DIRECTIONAL DYNAMICS CONNOTED BY VISUALLY PERCEIVED WORDS ON THE PHYSICAL POSITION OF THE APPARENT HORIZON

##### 1. Hypothesis

This experiment is concerned with the directional dynamics of stimuli of a different nature than those previously used. The hypothesis tested here deals with the question whether visually perceived words connoting directional dynamics by their meaning manifest effects on the position of the apparent horizon similar to those of the figures used in the first experiment. Our hypothesis is that the apparent horizon will shift relatively downward with words connoting upwardness as compared with words connoting downwardness.

##### 2. Procedure

The stimulus words, cut in the form of stencils with letters one inch high were placed in the light box so that the letters were dimly illuminated by a small bulb behind the flash opal glass screen. Thus, *S* saw the stimulus word in the form of stencil letters surrounded by darkness. The effect was one of a word "suspended in space."

regard we may point here to the fact brought out in a number of previous studies (1, 2, 4, 8) that generally, the apparent horizon is located below the objective horizon.



Ten words were employed; five words connoted upwardness: rising, climbing, raising, higher, upward; five words connoted downwardness: dropping, sinking, plunging, falling, downward.

In between trials *S* kept his eyes closed. Upon opening them he was required to fixate the center of the word, recite the word aloud and then to instruct *E* to move the word so that its center appeared at eye level.

Each of the 20 *Ss* was tested once with each of the 10 words. A 10 x 10 latin-square design was used with one man and one woman in each of the 10 sequences.

### 3. Results

The analysis of variance is presented in Table 2a; mean differences, and *t*-tests are presented in Table 2b. The *F*-test for differences due to dynamics in words is significant beyond the .05 level of confidence. The mean positions of the apparent horizon show that, with the exception of "raising," all of the up-words, as expected, show a greater shift of the apparent horizon in the downward direction than the down-words. With the exception of "raising," there is no overlap of the means of the up-words compared with the means of the down-words. Of the 25 comparisons (Table 2a) of the up-words with the down-words, 20 are in the expected direction.

As a further check on the overall effect of directional dynamics the data were reanalyzed as follows: For each *S* the mean position of the apparent horizon was computed for the five words connoting upwardness, and for the five words connoting downwardness. The means obtained were -2.0 cm. for the up-words and -1.2 cm. for the down-words. This difference is significant at the .05 level (one-tailed test).

TABLE 2a  
EXPERIMENT 2: EFFECT OF DIRECTIONAL DYNAMICS CONNOTED BY VISUALLY PERCEIVED  
WORDS ON THE PHYSICAL POSITION OF THE APPARENT HORIZON  
F-tests

Source of variation	df	Mean square	F	P
Subjects	19	224.55	—	—
Sequence	9	222.49	<1.00*	>.05
Sex	1	37.58	<1.00*	>.05
Subjects within sex and sequence	9	247.39	28.17	<.01
Conditions	9	17.84	2.03	<.05
Ordinality	9	17.61	2.00	<.05
Square uniqueness	72	11.94	1.36	>.05
Error	90	8.78		
Total	199			

\* Tested against subjects within sex and sequence.

TABLE 2b  
EXPERIMENT 2: EFFECT OF DIRECTIONAL DYNAMICS CONNOTED BY VISUALLY PERCEIVED WORDS ON THE PHYSICAL POSITION OF THE APPARENT HORIZON

Conditions	Means (cm.)	Means, Mean Differences, <i>t</i> -tests								
		2	3	4	5	6	7	8	9	10
1. Rising	-2.02	.77	2.15*	.68	.37	1.26	.38	1.04	1.15	.56
2. Climbing	-2.79		2.92**	.09	.40	2.03*	1.15	1.81	1.92*	1.33
3. Raising	+ .13			2.83**	2.52**	.89	1.77	1.11	1.00	1.59
4. Higher	-2.70				.31	1.94	1.06	1.82	1.83	1.24
5. Upward	-2.39					1.63	.75	1.41	1.52	.93
6. Dropping	-.76						.88	.22	.11	.70
7. Sinking	-1.64							.66	.77	.18
8. Plunging	-.98								.11	.48
9. Falling	-.87									.59
10. Downward	-1.46									

\* Significant at .05 level of confidence.

\*\* Significant at .01 level of confidence.

Though the overall means suggest that the apparent horizon shifts relatively opposite the dynamics connoted by the words, this conclusion is limited by the fact that there is variation in effect within each group of words (Table 2b). In particular, most noticeable, as mentioned above, is the reversal obtained for "raising." With regard to such differential efficacy of the single words it should be noted that these words were selected originally in an arbitrary manner by the experimenters, and at that time no attempt was made to assess the potency and quality of the directional dynamics of the words by a larger group of judges.

We have therefore extended the experiment to include such an assessment by a group of naïve judges. For this purpose a questionnaire was prepared and presented to 12 graduate students; they were asked to rate each word according to the intensity of their feeling about the direction conveyed by the word, on a continuous linear scale. A sample of the rating sheet is shown in Figure 2.

The results for the 12 Ss are shown in Table 3, which presents frequency of rankings in terms of intensity level as well as a score reflecting the

Two lists of words follow, one list contains words conveying an "up" direction; the other list contains words conveying a "down" direction. Please rank the words in each list according to the intensity of your feeling of the direction connoted by the words.

Place number of word at appropriate place on accompanying scale.

1. Climbing; 2. Rising; 3. Higher; 4. Upward; 5. Raising



Place number of word at appropriate place on accompanying scale.

1. Falling; 2. Sinking; 3. Plunging; 4. Dropping; 5. Downward



FIGURE 2  
SAMPLE OF RATING SHEET

TABLE 3  
FREQUENCY OF RANKINGS (12 JUDGES) IN TERMS OF INTENSITY LEVEL OF WORDS

	Rank					Weighted score*
	Most intense 1	2	3	4	Least intense 5	
Rising	5	3	1	3	0	26
Climbing	3	3	2	2	2	33
Raising	2	1	3	0	6	43
Higher	1	3	3	4	1	37
Upward	1	2	3	3	3	41
Dropping	0	3	3	3	3	42
Sinking	0	1	3	6	2	45
Plunging	8	1	0	0	3	25
Falling	2	6	3	1	0	27
Downward	2	2	3	2	3	38

\* The weighted score is determined by multiplying the number of judgments in each rank by the number of the rank. A weighted score of small magnitude, therefore, represents high intensity of the direction connoted by the word.

intensity level for the group as a whole. The two most intense up-words selected by the independent judges were "rising" and "climbing"; the two most intense down-words were "falling" and "plunging." It is interesting to note that "raising," which yielded the reversal in the perceptual test, is the only word rated as least intense within each word group by as large a proportion as 50 per cent of the judges.

In view of the striking differences of the words in regard to potency of dynamics, the two most extreme words of each word group were chosen for further comparisons with respect to their effect on the position of the apparent horizon. The mean position of the apparent horizon for "rising" and "climbing" was  $-2.4$  cm. while the mean position for "falling" and "plunging" was  $-.9$  cm. The mean difference,  $1.5$  cm., is significant ( $t = 2.46$ ; 19 df;  $p < .05$ ; one-tailed test).

Taking all the evidence together, the results indicate that up-down dynamics connoted by words shift the position of the apparent horizon opposite the direction of the dynamics. In particular, when words are selected, which according to the judgments of an independent group, possess relatively intense dynamics in opposite directions, statistically significant shifts of the apparent horizon are obtained in the direction predicted by the hypothesis.

#### D. DISCUSSION

In the first study of this series (9) we demonstrated the existence of dynamic properties of objects, usually referred to only phenomenologically,



by means of objective behavioral measurement. The present study extends the findings of the previous investigation in two ways. First, figural dynamics in the up-down direction has been shown here to have effects on space localization similar to those previously reported for the left-right direction. In particular, a figure with directional dynamics upwards has the effect of shifting the physical location of the apparent horizon relatively downwards compared with a figure with directional dynamics downwards. The second extension concerns the type of material used, namely, there is an efficacy of directional dynamics connoted by visually presented verbal symbols (words) which is similar to that inherent in pictorial forms.

As stated in the previous study (9) only a tentative interpretation of the observed effects of directional dynamics can be attempted. In that paper we have given such an interpretation within the framework of the sensory-tonic field theory of perception. Our general assumption has been that visual directional dynamics affects the state of the organism in a particular manner (see 9, pp. 63-4): we assume that pictorial forms with dynamics exert a "pull" on the organism in the direction of the dynamics which is counteracted by an organismic "pull" in the opposite direction<sup>4</sup> and further that the apparent horizon shifts in the direction of the counteractive force.<sup>5</sup>

This interpretation may be applied also to the visually perceived verbal symbols (words). The experiment with words was formulated on the assumption that certain words, such as those used here, may be perceived as objects with an inherent dynamics emerging from their meanings. If this assumption is accepted, then the visually presented words can be conceived to be in a class with objects or configurations possessing dynamic properties, e.g., pictures of pointing hands.<sup>6</sup> Thus, the interpretation given above can be considered to hold for both the findings with physiognomically perceived objects and with physiognomically perceived words.

---

<sup>4</sup>For the apparent horizon, just as previously stated for the apparent median plane, we leave open the question of whether such an organismic "pull" is restricted to the ocular-muscular system or involves a wider bodily distribution.

<sup>5</sup>The assumption of a shift of the apparent horizon in the direction opposite the counteractive force induced by dynamics is consistent with our interpretations of other findings where the principle of action-and-counteraction has been invoked (see 8, 10).

<sup>6</sup>We might note that Morris (5), whose point of view essentially is close to stimulus-response conceptualization, still recognizes that there are words, described here as physiognomic, which belong to the class of iconic signs; for him "... an iconic sign is any sign similar in some respects to what it denotes" (5, 191), e.g., onomatopoeic words. In our experiment "climbing" is an iconic sign if the subject perceives the upward quality when looking at the word. For definition and conceptualization of "physiognomic language" see Werner (7).

## E. SUMMARY

The present study, consisting of two experiments, is a continuation of previous work designed to demonstrate that visual directional dynamics exists as a behaviorally measurable event. The first of these experiments showed that a figure with directional dynamics upwards or downwards has the effect of shifting the physical location of the apparent horizon (eye line), opposite the direction of the dynamics. The second experiment used visually presented verbal symbols connoting upwardness or downwardness, e.g., "rising" vs. "falling." Analogous to the findings with meaningful pictorial material used in the first experiment, the physical position of the apparent horizon shifted opposite the direction of the dynamics conveyed by the words.

Thus, in addition to demonstrating that visual dynamics in the upward-downward direction can be measured behaviorally, the study shows that certain semantic aspects of words can be conceived and dealt with experimentally in a theoretical framework developed for problems of perception.

## REFERENCES

1. FITTS, P. M. A study of location discrimination ability. In P. M. Fitts (Ed.), *Psychological Research on Equipment Design*. Washington: U. S. Government Printing Office, 1947.
2. JAFFE, K. M. Effect of asymmetrical position and directional dynamics on configurations on the visual perception of the horizon. M.A. Thesis, Clark University, June, 1952.
3. KADEN, S. Effect of directional dynamics of configurations and visually perceived words on the position of the apparent horizon. M.A. Thesis, Clark University, June, 1953.
4. McDOUGALL, R. The subjective horizon. *Psychol. Rev. Mono. Suppl.*, 1903, 4, 145-166.
5. MORRIS, C. Signs, Language, and Behavior. New York: Prentice Hall, 1946.
6. WAPNER, S., WERNER, H., BRUELL, J. H., & GOLDSTEIN, A. G. Experiments on sensory-tonic field theory of perception: VII. Effect of asymmetrical extent and starting positions of figures on the visual apparent median plane. *J. Exp. Psychol.*, 1953, 46, 300-307.
7. WERNER, H. Comparative Psychology of Mental Development. Chicago: Follet, 1948.
8. WERNER, H., & WAPNER, S. Progress Report on the study of the basic mechanisms underlying perception-personality relationships, No. 4. National Institute of Mental Health, Research Project MH-348, February, 1954 (unpublished).
9. ———. Studies in physiognomic perception: I. Effect of configurational dynamics and meaning-induced sets on the position of the apparent median plane. *J. of Psychol.*, 1954, 38, 51-65.
10. ———. Toward a general theory of perception. *Psychol. Rev.*, 1952, 59, 324-338.

Department of Psychology  
Clark University  
Worcester 10, Massachusetts





phenomenological theory of individual behavior object to regarding the behavior as a "puppet, a mechanical recipient of stimuli" (11, p. 34). Mowrer's recent treatment of the subjects of neurosis, therapy, and learning (6) appears to be the most recent protest against a strictly mechanistic explanation of behavior.

Before we can profitably examine certain of Mowrer's views, a discussion of the nature of the concept of responsibility is indispensable. "That a person is legally responsible for a certain behavior . . . means that he is liable to a sanction (penalty) in case of contrary behavior" (2, p. 65). It is obviously not true that an individual can have knowledge of all the norms of a legal order, and therefore the contrary assumption is a legal fiction "against which no evidence is permitted" (2, p. 44). Primitive law is characterized by the principle of absolute responsibility, i.e., whether the accused "has anticipated or intended the effect of his conduct is irrelevant" (2, p. 65). Modern law has not completely abandoned this principle, since in some instances it attaches sanctions, though of a lesser degree of severity, to behavior by means of which a harmful effect has been brought about without the individual having intended or actually anticipated such a result (2). A further illustration of absolute responsibility pertains to sanctions which are directed against individuals who merely stand in a certain legal relationship to the individual who by his own conduct has committed an offense (2). Even in the case of the individual anticipating and maliciously intending a harmful effect, it must be recognized that all behavior is fully determined by the following conditions: the individual's past experiences, the present state of organism, and the present combination of external and internal stimuli impinging upon him; further, whatever action occurs under these conditions is the only one that can occur. The legal concept of responsibility requires as its correlate the fiction of a *free will* acting as an independent agent, which implies that more than one action may result under any set of these conditions. The arbitrary, normative connection between an action and the resulting liability to sanction is also evidenced by the fact that the "same" behavior at different phases of history does not always constitute an offense.

Thus it appears that responsibility is a concept of ethics and the legal law and not a concept of a natural science like psychology. This view is irreconcilable with the frequently held opinion that the term responsibility implies that a person elects a course of action out of several equally possible alternatives and that this course of action is a major determinant of subsequent events.

Mowrer accepts the deterministic point of view in stating that the behavior of organisms "is fully and lawfully determined" (6, p. 181). He



implicitly recognizes the normative-evaluative or legal nature of responsibility in affirming that "if the individual has done a socially objectionable act he has, by that very fact *acted irresponsibly*" (6, p. 181). His treatment of neurosis and psychotherapy, however, appears to be inconsistent with these premises and suggests an unrecognized tendency to establish a basis for "psychological" responsibility. He holds that "neurotics are *not* characteristically persons with deeply repressed libidinal and hostile impulses, but are rather the victims of their own devious efforts to make themselves immune to social responsibility" (6, p. 169).

Mowrer's views on anxiety provide the foundation for his treatment of neurosis and psychotherapy. He distinguishes between normal and neurotic anxiety. "Normal anxiety is usually of tolerable proportions" (3, p. 74) and is characterized as being "intelligible, (to the afflicted person, and) appropriate to the psychological situation" (3, p. 73). Neurotic anxiety, on the other hand, is perceived by the individual experiencing it as "meaningless, mysterious" (3, p. 77). "Anxiety in its neurotic sense can be understood only by positing . . . *repression*" (5, p. 281). Repression is referred to as "*mismanagement of normal anxiety*" (3, p. 74), "conscience-killing" (5, p. 284), and the "protective trick of lying" (6, p. 171). Repression, instituted "half deliberately, half automatically" (3, p. 74) is achieved by a "determined campaign of rationalization, forced 'forgetting', fault-finding in others" (5, p. 284). Mowrer believes that the distinction between these two types of anxiety is "essential to understanding the genesis of neurosis" (3, p. 73). The temporary relief gained through repression is not "a dependable thing and, when repression fails . . . the terror of neurosis is the outcome" (5, p. 284). The sequence of events leading to neurosis as conceived by Mowrer thus appears to be as follows: certain behavior arouses the superego which leads to normal, intelligible anxiety. Mismanagement of such normal anxiety (repression) frequently achieves only temporary relief and the "return of the repressed" (5, p. 284) is accompanied by neurotic, mysterious anxiety (neurosis).

In general then, Mowrer holds that neurosis involves a "fission between the id-ego and the superego rather than between the ego-superego and id" (3, p. 89). The latter constellation he considers the traditional view of psychoanalysis and also that of Dollard and Miller (1) in their treatment of neurosis within the conceptual framework of reinforcement theory.

For purposes of brevity, the following discussion will be chiefly limited to questioning the methodological tenability of Mowrer's position. His distinction between normal and neurotic anxiety is based upon the criteria of understanding, appropriateness, and management. These terms, as well

## REFERENCES

1. DOLLARD, J., & MILLER, N. E. *Personality and Psychotherapy*. New York: McGraw-Hill, 1950.
2. KELSEN, H. *General Theory of Law and State*. Cambridge: Harvard Univ. Press, 1949.
3. MOWRER, O. H. Neurosis and psychotherapy as interpersonal processes: a synopsis. In Mowrer, O. H., *Psychotherapy, Theory, and Research*. New York: Ronald Press, 1953.
4. ———. Neurosis, psychotherapy, and two-factor learning theory. In Mowrer, O. H., *Psychotherapy, Theory, and Research*. New York: Ronald Press, 1953.
5. ———. Neurosis: a disorder of conditioning or problem solving? *Ann. New York Acad. Sci.*, 1953, **56**, 273-288.
6. ———. Motivation and neurosis. In *Current Theory and Research in Motivation*. Lincoln: Univ. Nebraska Press, 1953.
7. MURPHY, G. *Historical Introduction to Modern Psychology*. New York: Harcourt, Brace, 1950.
8. NEWCOMB, T. M. Comments on papers by Professors Nowlis and Mowrer. In *Current Theory and Research in Motivation*. Lincoln: Univ. Nebraska Press, 1953.
9. NOWLIS, V. Comments on papers by Professors Newcomb and Mowrer. In *Current Theory and Research in Motivation*. Lincoln: Univ. Nebraska Press, 1953.
10. SHOBEN, E. J., JR. Some observations on psychotherapy and the learning process. In Mowrer, O. H., *Psychotherapy, Theory, and Research*. New York: Ronald Press, 1953.
11. SNYGG, D., & COMBS, A. W. *Individual Behavior*. New York: Harper, 1949.
12. TOULMAN, E. C. Cognitive maps in rats and men. *Psychol. Rev.*, 1948, **55**, 189-208.

4409 West Point Loma Boulevard  
San Diego 7, California

## THE INFLUENCE OF HYPOTHALAMIC STIMULATION ON EVOKED CORTICAL POTENTIALS\*<sup>1</sup>

*Laboratory of Neurophysiology, University of Minnesota Medical School*

E. GELLHORN, W. P. KOELLA, AND H. M. BALLIN

### A. PROBLEM

It was shown in a preceding investigation (2) that nociceptive and acoustic or optic impulses may interact, as indicated by the increased response of a cortical projection area to specific stimuli during mild nociceptive stimulation. Nociceptive impulses act on the cerebral cortex as a whole via the hypothalamic-cortical system. This is suggested by the fact that the generalized cortical reaction is associated with excitation of the hypothalamus (3, 1) and is ipsilaterally diminished or abolished by unilateral hypothalamic lesions (6). In view of the importance of these findings for the problem of consciousness (4), it was deemed advisable to study this interaction directly by determining the influence of hypothalamic stimulation on the response of specific sensory projection areas to acoustic or optic stimuli.

### B. METHODS

The experiments were performed on nine cats anesthetized with Dialurethane or nembutal. The methods used for exposure of the cerebral cortex, recording of action potentials, and acoustic and optic stimulation were as described in earlier papers. The hypothalamus was stimulated by means of a square wave generator. In a large number of the experiments a device of O. Schmitt (10) was used in order to eliminate stimulation artifacts.<sup>2</sup> The site of stimulation was determined in serial sections by the use of Hess' prussian blue method.

### C. RESULTS

#### 1. *The Influence of Hypothalamic Stimuli on the Response of the Auditory Projection Area to Acoustic Stimuli*

Figure 1A shows on a strychninized site of the auditory projection area

\*Received in the Editorial Office on August 7, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This work was carried out with the aid of a grant from the Louis W. and Maud Hill Family Foundation.

<sup>2</sup>A much greater voltage has to be applied when this device is used than when it is not.

that without stimulation the spikes appear mostly in groups of two and three, whereas during hypothalamic stimulation the grouped discharges gradually disappear and are replaced by more continuous discharges which occur at somewhat irregular intervals. On cessation of this stimulus the control pat-

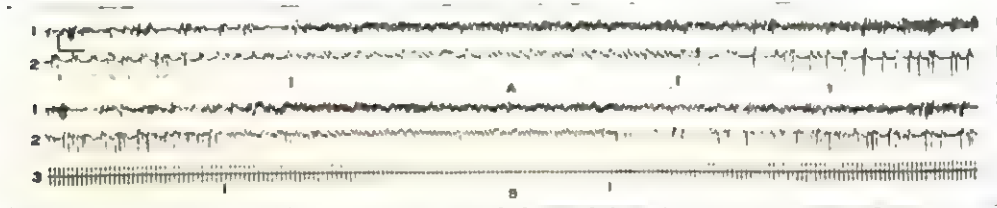


FIGURE 1

Cat (nembutal). (1) posterior sigmoid gyrus; (2) ectosylvian gyrus (auditory area, strychninized); (3) acoustic stimulus. Between the vertical lines (below Part A and B respectively) stimulation of the hypothalamus (in upper posterior infundibular region on right side) involving Schmitt's device (+0V, 110/sec; pulse duration 0.2ms). Calibration: 400  $\mu$ V; 5 sec.

tern reappears. The sensori-motor area shows under conditions of hypothalamic stimulation the loss of "Dial" potentials and the increased frequency and amplitude of background potentials which was described earlier (7). The repetition of this experiment under conditions of acoustic stimulation (Figure 1B) shows the following changes:

1. The spike frequency in the auditory area is greater during acoustic stimulation than in the absence of these stimuli. Although some spikes appear to be related to the acoustic stimuli, others are not and frequently gaps in the auditory response are noted.

2. During the period of hypothalamic stimulation the spikes appear to an increasing degree in response to the acoustic stimuli. This effect persists for some time after the cessation of hypothalamic stimulation.

Results of this kind have been obtained more than 50 times in nine cats. The following experiments are reported in order to illustrate more specifically the conditions of this summation phenomenon.

A comparison of Figure 1 with Figure 2 (obtained within a few minutes from the same animal) shows that the summation phenomenon occurs regardless of whether the hypothalamic stimulation per se results in an increase (Figure 1A) or a decrease in the spike frequency<sup>3</sup> of the cerebral auditory area (Figure 2B). The latter experiment involved a somewhat

<sup>3</sup>Hypothalamic stimulation may decrease or increase spike frequency (Murphy and Gellhorn, 1945; Gellhorn and Ballin, 1948) depending on the intensity of hypothalamic stimulation and the degree of excitation of the cortical area.



lesser degree of strychninization and a lesser intensity of hypothalamic stimulation. Summation occurs, but the effect is less than in Figure 1, as noted particularly in the lessened duration of the summation following the period of stimulation.<sup>4</sup>

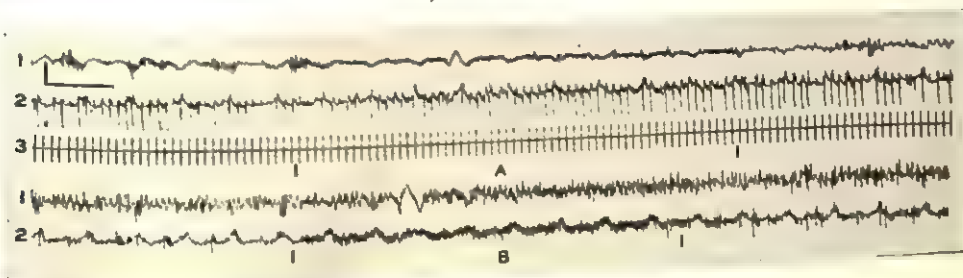


FIGURE 2

Same cat, same sites of recording and stimulation as in Figure 1. Hypothalamic stimulation at 30V. Calibration: 400  $\mu$ V; 10 sec.

Although in general weak hypothalamic stimuli reduce the frequency of cortical spikes whereas strong stimuli increase it, the selection of an adequate hypothalamic stimulus as such is no guarantee of the demonstrability of the summation phenomenon since a suitable form of acoustic (or optic) stimulation is likewise necessary. If the acoustic stimulus is too weak or too frequent, summation fails to appear. Frequently it was found practical to begin the experiment with an acoustic stimulus which elicited a 100 per cent response and then to reduce the intensity or increase the frequency of this stimulus until the response became irregular. Then hypothalamic stimulation often induced the summation phenomenon and frequently the response became 100 per cent. In other experiments the summation phenomenon was studied in auditory fatigue. For this purpose acoustic stimulation which evoked an auditory potential in response to each stimulus was continued for several minutes until the percentage of evoked potentials was greatly reduced. Under these conditions hypothalamic stimulation augmented the auditory response considerably and frequently reestablished a 1:1 relation between stimulus and effect.

Occasionally very striking summation phenomena are observed. Figure 3 illustrates an experiment in which the auditory response was practically absent under control conditions (the dots indicate the frequency of the acoustic stimulus<sup>5</sup> and allow one to identify a trace of a cortical response

<sup>4</sup>Subsequently referred to as "afterdischarge."

<sup>5</sup>The acoustic stimulus is applied at this frequency throughout the period (before, during, and after hypothalamic stimulation) shown in Figure 3.

between the grouped potentials). Under hypothalamic stimulation a regular response develops and continues for several seconds after the cessation of the hypothalamic stimulus.

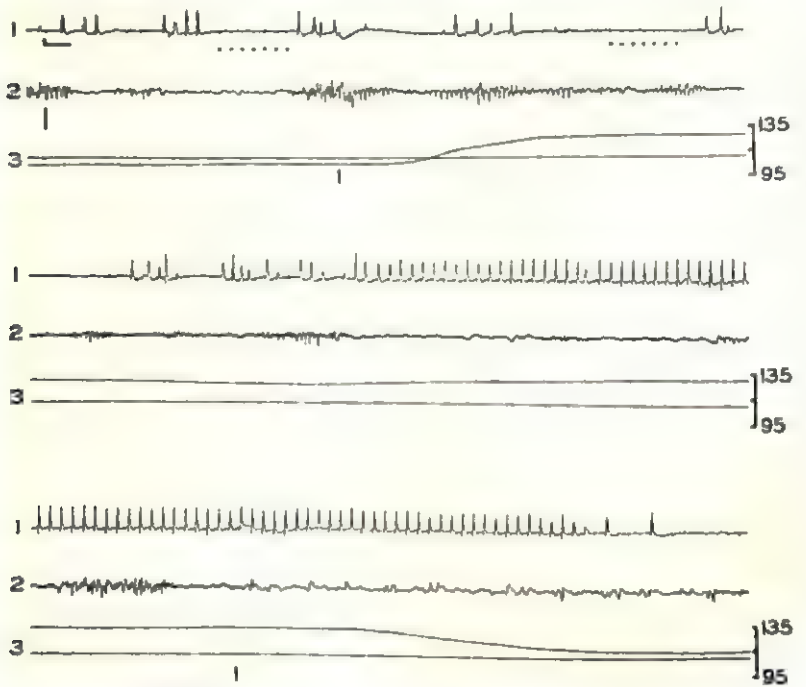


FIGURE 3

Cat (Dial-urethane). (1) ectosylvian gyrus (auditory area), strychninized; (2) posterior sigmoid gyrus; (3) blood pressure. Left upper posterior hypothalamus stimulated at 4.0 V, 51/sec between vertical lines. Calibration 400  $\mu$ V; 1 sec. Continuous record. Details in text.

In a group of experiments intensity and frequency of hypothalamic stimulation were varied. It was found that frequencies below 40/second were usually ineffective.

Within the effective range an increase in frequency and/or intensity augmented the summation phenomenon by a reduction in summation time and lengthening of the afterdischarge. The summation phenomenon showed a long summation time, often 10-15 seconds, although the excitatory effect on the cortical potentials (disappearance of grouped potentials, etc.) occurred almost immediately (see Figures 1, 2, 4, 5). In some cases the summation phenomenon occurred while the blood pressure and pulse rate were increased (Figure 3). However, after appropriate changes in the parameters

of stimulation hypothalamic excitation did not alter the blood pressure and yet the summation phenomenon persisted (Figures 4 and 5). In other experiments in which the summation time was very long or in which summation appeared only after the stimulation period, the summation occurred as the blood pressure fell slightly below the control level. It may be concluded that this summation is not dependent on changes in the blood pressure.

If the hypothalamus is stimulated while potentials are recorded from the contralateral hypothalamus and the cortex, it can be shown that the summation phenomenon is associated with electrical signs of hypothalamic excitation. This is illustrated in Figures 4 and 5 on spikes which appeared in the hypothalamus as the result of repeated application of strychnine to the cortex (7). Stimulation of the left posterior infundibular region of the hypothalamus led to degrouing of spikes in the contralateral hypothalamus

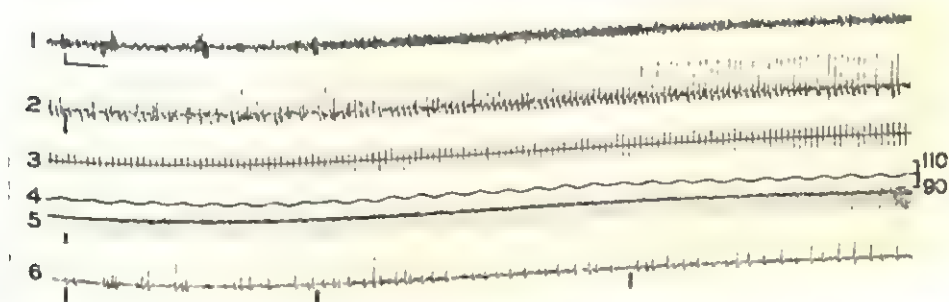


FIGURE 4

Cat (nembutal). (1) posterior sigmoid gyrus; (2) ectosylvian gyrus (auditory area, strychninized); (3) acoustic stimulus, (4) blood pressure; (5) pulse rate; (6) right hypothalamus. Stimulation of the left hypothalamus (infundibular area) involving Schmitt's device at 35V, 41/sec between vertical lines. Calibration: 500  $\mu$ V; 5 sec.

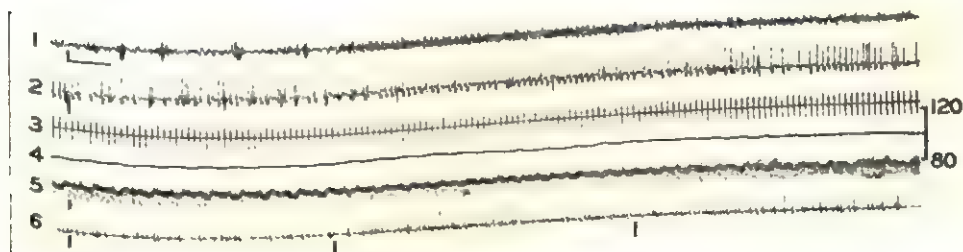


FIGURE 5

Same cat, same sites. Stimulation of hypothalamus at 40V, 110/sec. Calibration 500  $\mu$ V; 5 sec.

(Figure 4) or to an increase in the frequency and amplitude of the spikes and an augmentation in the background potentials (Figure 5).

Thus far the criterion of hypothalamic and acoustic summation has consisted in the increase of the frequency at which the cortex responded to acoustic stimuli during or immediately after stimulation of the hypothalamus. However, it was also noted that the amplitude of the auditory potentials may be increased under these conditions. Furthermore, the cortical auditory response may become more complex and prolonged (multiple spikes, Figure 6, channel 2). A more detailed study of these phenomena by means of oscillographic record may be given later.

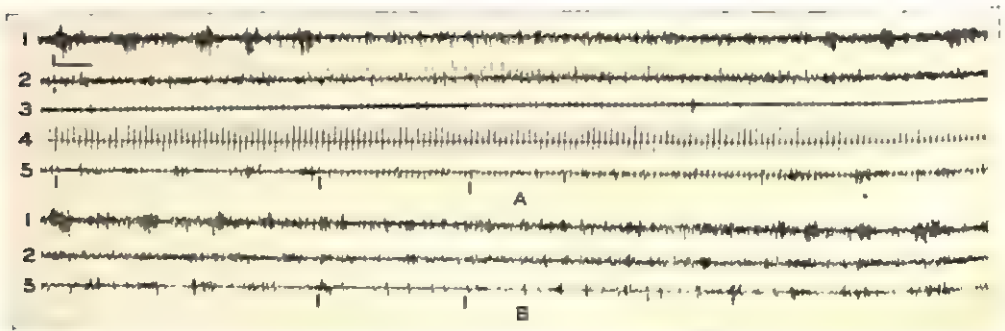


FIGURE 6

Cat (nembotal). (1) and (2) in left auditory area, several millimeters apart; (3) acoustic stimulus; (4) optic stimulus; (5) visual projection area. Hypothalamic stimulation (lateral and just anterior to right mammillary body) involving Schmitt's device at 25V, 110/sec, pulse duration 0.2 ms. Fig. 6A with acoustic and optic stimuli; Fig. 6B without them. Calibration 200  $\mu$ V; 5 sec.

The interaction between the hypothalamic and acoustic stimuli on the auditory system may be demonstrated in still another manner. As mentioned previously, the effect of hypothalamic stimulation on normal cortical potentials (no application of a convulsant to the cortex) consists in the elimination of grouped ("Dial") potentials and their replacement by frequent background potentials. This effect on the auditory projection area is greatly intensified if acoustic stimuli are applied. Figure 6 shows that the duration of this period of excitation is greatly prolonged in the presence of acoustic stimuli. (Cf. the records of No. 1 in Figures 6A and 6B.) It also shows that in response to the acoustic stimuli multiple spikes appear during the hypothalamic activation which are absent in the preceding and following control period (Record 2 in Figure 6A).

Figure 7 illustrates an experiment in which strychnine spikes are recorded from the auditory and visual projection areas. Under the influence of



hypothalamic stimulation a degrouping of the potentials occurs which is sharply confined to the period of stimulation (Figure 7B). However, if the hypothalamus is excited during acoustic stimulation the summation phenomenon occurs and is extended into the post-stimulatory control period.

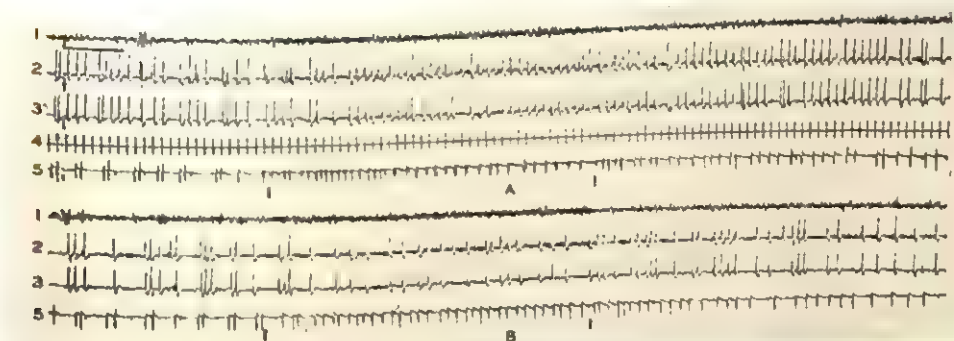


FIGURE 7

Cat (nembutal). (1) posterior sigmoid gyrus; (2) and (3) auditory projection area, strychninized; 4, acoustic stimulus; (5) visual projection area. Stimulation (Schmitt's device) in left infundibular area at 35V, 110/sec, pulse duration 0.2 ms. Figure 7a with acoustic stimuli; Figure 7b without them. Calibration 500  $\mu$ V; 5 sec.

In addition, some unspecific summation effect on the visual projection area seems to result from acoustic plus hypothalamic stimuli which appears in the form of double spikes (No. 5, Figure 7A). The latter are not seen when the hypothalamus is stimulated in the absence of acoustic stimuli (No. 5, Figure 7B). In contradistinction to the long persistence of the summation phenomenon in the auditory area, the hypothalamic stimulation on the optic area remains confined to the period of hypothalamic stimulation. It seems to follow that the hypothalamic stimulus interacts with the specific sensory stimulus chiefly on the specific sensory projection area; but to a lesser extent a specific sensory stimulus (acoustic) may increase the response of another projection area (visual) to the hypothalamic stimulus.

## 2. *The Interaction of Optic and Hypothalamic Stimuli on the Visual Projection Area of the Cortex*

Similar experiments were conducted on the action of simultaneous optic and hypothalamic stimulation on the visual projection area in which normal potentials or strychnine spikes served as indicators.

Figure 8B shows that hypothalamic stimulation induces an asynchrony (disappearance of grouped "Dial" potentials) whereas the same stimuli when

applied in conjunction with optic stimulation cause a much greater excitation of the cortex (Figure 8A). The latter appears not only in the form of asynchrony but also as recruitment of additional neurons, indicated by the increased frequency and amplitude of "background potentials." The

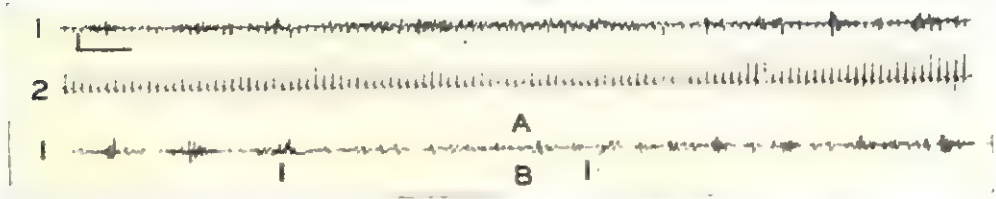


FIGURE 8

Cat (nembotal), site of stimulation as in Figure 6. (1) visual projection area; (2) optic stimulus. Hypothalamic stimulation (with optic stimulation in Figure 8a, but without it in Fig. 8b) involving Schmitt's device at 30V, 110/sec, pulse duration 0.1 ms. Calibration: 200  $\mu$ V; 4 sec.

amplitude of the optic response seems likewise increased<sup>6</sup> during the first half of the stimulation period before the optic potentials become submerged in the "recruited" background potentials.

Still clearer effects are obtained after strychninization of a site of the visual projection area. Figures 9 and 10 show the effect of hypothalamic and optic stimulation at varying degrees of cortical strychninization. In Figure 9, while the action of strychnine is still weak, only a few spikes

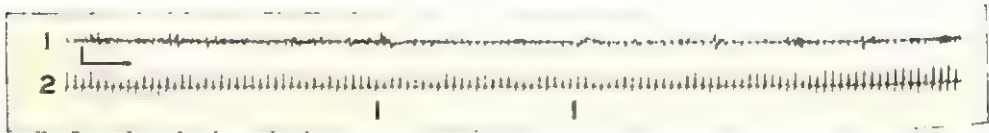


FIGURE 9

Same cat as in Figure 8. (1) visual projection area (strychninized); (2) optic stimulus. Hypothalamic stimulation as in Fig. 8. Calibration 200  $\mu$ V; 5 sec.

appear during the control period and the optic response is small and irregular. During hypothalamic plus optic stimulation the strychnine spikes disappear, but the optic response becomes distinct and regular, and this effect continues for some time after stimulation. A few minutes later the strychnine effect has become much stronger, as indicated by the amplitude and frequency of the spikes. However, the spikes are irregular and do not show any consistent relation to the optic stimuli (Figure 10a). During

<sup>6</sup>See also the optic response in Figure 6a.

hypothalamic stimulation the optic stimuli become effective. A 1:1 relation between optic stimuli and strychnine potentials persists throughout the whole stimulation period and for several seconds afterwards. Finally two experiments are presented in Figures 10b and 10c on the same animal in which

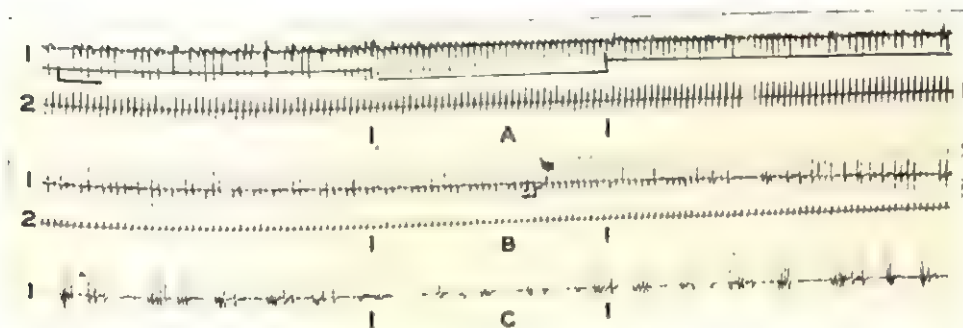


FIGURE 10

Same cat, sites and stimulation as in Figure 9. (1) visual projection area; (2) optic stimulus. Figure 10a & b. Hypothalamic plus optic stimulation. Note the intensity of the optic stimuli is reduced in Figure 10b. Figure 10c. Hypothalamic stimulation without optic stimuli. Calibration 200  $\mu$ V; 5 sec.

the effect of hypothalamic stimulation with and without optical stimulation may be compared. It is noted that whereas hypothalamic stimulation per se reduces the number of spikes (Figure 10c) a marked increase in their frequency occurs during hypothalamic plus optic stimulation. A close inspection reveals that the small visual potentials are converted into convulsive spikes during hypothalamic stimulation (Figure 10b). That hypothalamic stimulation increases the response of the cortex to sensory stimulation is suggested by the fact that an increase in the convulsive response to optic stimuli similar to that accomplished by hypothalamic stimulation occurs if, in the absence of the latter, the intensity of the optic stimuli is increased.

#### D. DISCUSSION

The experiments show that hypothalamic stimulation facilitates the response of the cortex to acoustic and optic stimuli. This summation phenomenon may be demonstrated through normal as well as convulsive potentials. It may appear in several forms: (a) Near-threshold stimuli (acoustic or optic) elicit a cortical response more often during hypothalamic stimulation than under control conditions. (b) Even subthreshold stimuli may become effective. This is seen particularly in experiments involving strychninization of the projection area. (c) The evoked potentials may indicate

multiple discharges under the combined influence of hypothalamic and sensory stimulation, although such an effect is absent if either stimulation is applied alone. (*d*) If the cortex is strychninized lightly so that optic or acoustic stimuli evoke a convulsive potential only in response to a fraction of the stimuli, this percentage will be greatly increased when the hypothalamus is stimulated at the same time.

Since hypothalamic stimulation induces an excitation of the whole cortex, it may be assumed that the result of such summation processes will likewise appear in other cortical projection areas studied under similar circumstances.

To what extent the interaction takes place at the thalamic and at the cortical level has not yet been determined. However, it may be safely assumed that a neurogenic interaction, and not circulatory changes, is at the base of the potential changes described in this paper. These are the reasons: (*a*) Striking examples of summation may occur although the blood pressure remains unchanged; (*b*) in experiments in which the blood pressure rises, the increased cortical responsiveness may occur a considerable time after the maximum of the pressor effect, and it may continue while the pressure falls to or slightly below the control level; (*c*) slight changes in intensity or frequency of sensory (acoustic or optic) stimulation may abolish the summation although the pressor effect remains unaltered; (*d*) increasing intensity of hypothalamic stimulation may increase the summation phenomenon and its duration (shortening the summation time and increasing the after effect) in the absence of a pressor effect.

It is interesting to note that the increased cortical excitation which results from the interaction between sensory impulses and hypothalamic stimuli is not restricted to the specific projection area. Although summation processes are most pronounced in the auditory projection area, if acoustic and hypothalamic stimuli are applied at the same time, evidence has been presented to show that this combination of stimuli may exert a greater excitatory effect on the visual projection area than either form of stimulation when presented alone.

The summation phenomenon described in this paper seems to be based on the interaction of the effects of hypothalamic excitation (shown in the contralateral hypothalamus, Figure 4) with those induced by the activation of the classical afferent systems. Bearing in mind the fact that the activation of the latter systems is effective in evoking potentials in deep anesthesia in which sensations or perceptions are absent, these potentials are obviously not the physiological correlates of conscious sensations. On the other hand,



the relation of hypothalamic-cortical discharges to awareness is well established by the work of Bernhaut *et al.*, and that of Magoun and collaborators (5). It is suggested that the interaction of hypothalamic and sensory discharges represents minimal requirements to induce sensations and perceptions. It remains to be seen whether still other factors are necessary for these processes.

### E. SUMMARY

The study indicates that the interaction between hypothalamic stimuli and acoustic or optic stimuli on the auditory or visual cortical projection area respectively leads to summation phenomena. They are characterized by:

1. An increased frequency of the cortical response to the sensory impulses provided that not every impulse elicits a cortical response under control conditions.

2. An increase or, if previously not present, an appearance of multiple sensory discharges during hypothalamic stimulation.

3. Unspecific effects which occur outside the specific projection area. Thus hypothalamic plus acoustic stimuli cause a more prolonged disappearance of "Dial"-potentials in non-auditory areas than is induced by hypothalamic stimulation per se.

This interaction occurs frequently with a summation time of 10-15 seconds or more. The effect may persist for varying periods of time following the cessation of hypothalamic stimulation (afterdischarge). The summation is not caused by circulatory changes since it is found whether the blood pressure is unchanged, rises, or falls slightly. The summation takes place regardless of whether the stimulation of the hypothalamus (without sensory stimuli) caused a decrease or an increase in cortical potentials.

The summation phenomenon has been shown to occur for optically or acoustically induced normal as well as convulsive potentials (induced by cortical application of strychnine).

With increasing intensity and/or frequency of hypothalamic stimulation the summation occurs earlier and shows increasing degrees of afterdischarge.

It is suggested that the interaction between sensory and hypothalamic impulses which leads to the summation phenomenon plays an important part in the production of conscious sensation and perception.

### REFERENCES

1. BERNHAUT, M., GELLHORN, E., & RASMUSSEN, A. T. Experimental contributions to problems of consciousness. *J. Neurophysiol.*, 1953, **16**, 21-35.
2. GELLHORN, E., KOELLA, W. P., & BALLIN, H. M. The interaction between acous-

- tic or optic impulses and nociceptive impulses on the cerebral cortex. *J. Neurophysiol.*, 1954, 17, 14-21.
3. GELLHORN, E., & BALLIN, H. M. The effect of afferent impulses on hypothalamic potentials. *Amer. J. Physiol.*, 1946, 146, 630-635.
  4. GELLHORN, E. *Physiological Foundations of Neurology and Psychiatry*. Minneapolis: Univ. Minnesota Press, 1953.
  5. LINDSLEY, D. B., BOWDEN, J. W., & MAGOUN, H. W. Behavioral and E.E.G. changes following chronic brain stem lesions. *EEG Clin. Neurophysiol.*, 1950, 2, 483-498.
  6. KOELLA, W. P., & GELLHORN, E. The influence of diencephalic lesions upon the action of nociceptive impulses and hypercapnia on the electrical activity of the cat's brain. *J. Comp. Neurol.*, 1954, 100, 243-256.
  7. MURPHY, J. P., & GELLHORN, E. The influence of hypothalamic stimulation on cortically induced movements and on action potentials of the cortex. *J. Neurophysiol.*, 1945, 8, 341-364.
  8. ———. Further investigations on diencephalic-cortical relations and their significance for the problem of emotion. *J. Neurophysiol.*, 1945, 8, 431-448.
  9. RHINES, R., & MAGOUN, H. W. Brain stem facilitation of cortical motor response. *J. Neurophysiol.*, 1946, 9, 219-229.
  10. SCHMITT, O. Tissue stimulators using radiofrequency coupling. *Rev. Sci. Instru.*, 1949, 20, 170-173.

117 Owre Hall

Medical School

University of Minnesota

Minneapolis 14, Minnesota

## A MEASURE OF SOCIAL CONFORMITY\*<sup>1</sup>

*Department of Psychology, Los Angeles State College*

---

RAYMOND E. BERNBERG

---

### A. THE PROBLEM

There exist in personality development certain influences of culture and society which are restrictive in nature (4, p. 156). The conformity to cultural and social pressures becomes manifest in the acts and attitudes of members of a society. The concept of superego is a well known expression which infers such social restrictive behavior. Social conformity is defined by the tendencies of members of a society to manifest communality of attitudes and of behavior as a result of these influences.

The purpose of the research reported here is to measure such conformity as it is manifested in attitudes and to test the hypothesis that such measurements can be used to predict behavior. If the measure is capable of predicting differences in groups and individuals its utility as a screening and diagnostic device in measuring social conformity may well be considered in a variety of situations.

For this purpose, social conformity has been defined in terms of a number of determinant areas, and an attitude scale has been constructed of items which sample the areas.

The determinants were outlined as follows:

(a) *Moral values* as manifested in attitudes of responsibility toward groups, through typical sexual attitudes, through attitude toward law, government, etc.

(b) *Positive goals* as manifested through attitudes toward long range planning, time perspective, through consistent attitudes toward shifting goals.

(c) *Reality testing* as manifested in awareness of others' attitudes toward him, learning by experience, the projection of reality to any life role.

(d) *Ability to give affection* as manifested in attitudes toward marriage, family, children, attitude toward perseverative relationships, attitude toward women and sexual relationships.

---

\*Received in the Editorial Office on August 13, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>Based on a paper presented at the meeting of the Western Psychological Association on April 25-26, 1952.

(e) *Tension level* as manifested in attitude about concern with intimates, empathy and identification, attitude toward personal threat, degree of self-satisfaction.

(f) *Impulsivity* as manifested in lack of inhibitive attitude patterns.

## B. CONSTRUCTION OF THE SOCIAL CONFORMITY SCALE

The attitude scale has utilized a process of indirect or projective measurement which has been described elsewhere (1, 2) as the direction of perception technique of attitude measurement. In this technique, the items are of multiple choice form and ostensibly call for a judgment in quantitative terms about a presumably determinative fact. Actually there is no known foundation in fact for a choice among the possible answers. Hence the tendency of the subject to deviate toward one extreme or the other is presumed to express a "direction of perception" based on his need-value system.

*For example:* (The following item reflects the third determinant.)

"Public opinion polls show what percentage of people feel it is silly to make close friendships because few people can really understand you?"

a) 30% b) 40% c) 50% d) 60% e) 70%

The manner in which a score for "conformity" is arrived at will be set forth later.

In this report, results will be presented based upon two types of inter-group comparison. First, the use of the technique of attitude measurement, "direction of perception" in this study will be justified by showing that the scores are not significantly affected by educational levels, intelligence, socio-economic level, religion, age, sex, nor cultural background in our society. Secondly, it will be shown that significant differences do occur between a normal population and criterion groups of individuals manifesting socially deviant behavior.

Sixty-eight questions were constructed with approximately equal distribution of the determinants. Each question had five choices for answer. The choices were numerical percentages with equal numerical distances between adjacent choices. Two forms of the test were constructed. While each form contained the same questions, the range of numerical percentages in the choices for answers was different. In Form A, the range of values on a percentage scale began fairly high as a rule. The lowest value for a choice of answer on Form A was the highest value for a choice of answer on Form B. Form B contained a range of values for the choices of answer



which began fairly low and the highest value for a choice of answer was the lowest for Form *A*.

The following is an example:

"Statistics show what percentage of men like to write things on the walls in men's rooms?"

Form *A*: a) 27% b) 40% c) 53% d) 66% e) 79%

Form *B*: a) 11% b) 40% c) 19% d) 23% e) 27%

The scale was entitled *Human Relations Inventory*. The instructions on the test sheets gave no information as to the purpose of the test. In general, the whole situation was non-revealing.

### C. THE PRELIMINARY EXPERIMENT

The *HRI* (Human Relations Inventory) was administered to two groups. One included 100 graduating senior boys of median and modal age, 18, who were in a high school of extremely low socio-economic level, high degree of Mexican national origin as well as negro racial background, and whose mean *IQ* on the Otis short form scale is 92. The other group of 125 was a comparable age group from a different high school but whose mean *IQ* was 114 and who come from a broad range of socio-economic level with a preponderance in the middle and upper levels. In addition, this population was entirely caucasian. Approximately one-half of each group took each Form *A* and *B*, giving four groups in all.

### D. RESULTS AND DISCUSSION

The forms were analyzed by a tabulation of the distribution of responses in the choices of answer for each item in each form for all groups. Form *A* and *B* were compared by analysis of the responses to the different range of values for each form. Allport's index of institutionalization (3, p. 82-83) was computed for the distribution of responses to the five choices of answer for each item. This was to determine the degree of conformity of response to the items. It indicates the fact of chance distribution over the five choices of answer or the degree to which the responses take on a *J* curve of conforming distribution with the mode on one end of the five choices of answer. In other words, this index indicates the relative communal meaningfulness each question had for the subjects.

An index of 2.00 indicates a chance distribution in the frequency of responses to the five choices for answer. An index of 3.00 indicates a normal distribution in the frequency of responses to the five choices for answer with

the mean at the mode. This indicates a conforming distribution. An increasing index in numerical value from three upward indicates relatively more intense conforming distributions.

A similarity of conformity in response to the same item in both forms would be an index of the operation of direction of perception. The direction of perception would indicate a personal frame of reference channeling the response toward "more" or "less." This would be an outcome of the operation of the social conformity process.

The correlation of the index of institutionalization for the same item, on the two forms, was found to be .66 ( $P < .001$ ). The mean index for each form was computed and a test of significance made between the two means. For Form *A*—with a high range of values in the choices for answer—there was a mean index of 2.75, with a range from 1.9 to 11.5. The mean for Form *B* was 3.3 with a range from 1.8 to 7.3. The distributions of the two forms are not significantly different except for the case of one higher index in Form *A*.

We can now answer a fundamental question: "Does a comparison of results from the two forms indicate that the direction of perception is a stronger determinant of response than the absolute values of the choices offered?" Yes, because the high correlation indicates that although there might be slight mean differences in the two distributions, the systematic similar results in the responses to the same item with different ranges of values for the choices of answer reflects the operation of a direction of perception in the responses of the subjects to the various items.

A second question is in order. "Do socio-economic or intellectual factors obscure the operation of direction of perception?" The answer is "no," because both groups in taking both forms came out with almost identical results in the distributions of the frequency of responses to the choices for answer. This is significant in the establishment of the communal meaningfulness of the items to groups with broad differences in intelligence and social class. A measure of social conformity must necessarily be devoid of contamination with any such covariates.

Out of the 68 questions used in the inventory, 37 had an index of more than an approximate three, which makes them useful as indirect attitude items. These items were given weights dependent upon their degree of conformity in response in the standard population. In this manner, intensity is determined empirically. If we desire to measure non-conformity, then the relative modality of response in a specified standard population must be negated and increasing intensity of score is to be given to the choices of

answer receding in frequency of response from the modal extreme. Form *A* responses were taken and scored according to this method of measuring intensity of response. They provide a normative basis for comparison with a non-conformity group.

### E. THE VALIDATION EXPERIMENT

The next step was to administer the revised *HRI* to an acceptable non-conforming group. The subjects used here were 160 male inmates of a California youth prison, median and modal age of 21 years. This population is composed of offenders of the law concerning sex offenses, larceny, and assault. They constitute a behaviorally defined non-conforming group.

A comparison of mean scores reveals the standard population mean 9.8 (*SD* 4.4) and the non-conforming group mean is 12.3, (*SD* 5.8). This is a significant difference,  $p > .001$ , and in the predicted direction.

The items up to this point had been weighted based upon the empirical distribution of responses on them by the standard population. The operation of a difference in the right direction between these two populations suggested possible further weighting of the items based upon the empirical distribution of responses on them by the non-conforming group. These latter responses were analyzed yielding 17 of the 37 items whereby the non-conforming group indicates a less intense conforming distribution of responses on the items than the standard population.

The operation of this difference was maximized by the additional weighting of the 17 items. The *HRIs* of the non-conforming group were rescored to compare with a new sample of a standard population based upon the new weighting scheme. The sample selected for this was 108 male and female college students whose mean age is 25.7 years. Previous to this an analysis was made on 100 male and female college student *HRIs* as to possible differences between the sexes and whether age produced any differences. The results here were specifically negative illustrating the additional facts that age, up to 35 years, and sex could not be differentiated by this scale. In addition, a breakdown by religion was analyzed using these 100 cases plus the other 108. The results were completely negative.

A comparison of mean scores on the new weighted scale reveals the standard population mean 12.5, sigma is 6.9, and the non-conformity group mean is 18.6, sigma is 5.4. This is a significant difference,  $p > .001$ .

To add to the apparent validity of this scale, an homogeneous group was needed who by interest and intent could be designated a conformity group. The sample selected for this are 101 Police Science major students attend-

ing colleges in Los Angeles. Their mean score is 11.4, the sigma is 6.0, and a test of significance between the Police Science students and the prison group yields a significant difference,  $p > .001$ .

#### F. ADDITIONAL VALIDATION DATA AND RESULTS

The preliminary results of the social conformity scale indicated further study to establish (a) reliability, (b) cross-validation, and (c) additional validity measures.

A reliability coefficient of .77 was obtained with the Spearman-Brown prophecy formula using the youth prison group. Since the 37 items of the test have different weights, a random selection of heavily weighted and less weighted items were obtained for approximately two equal halves of the test.

Considering cross-validation, additional populations were sought for both standard and deviant groups. The quartile data presented in Table 1 indicate the ability of the test to replicate results and indicate successful cross-validation.

TABLE 1  
QUARTILE SCORES FOR THE SOCIAL CONFORMITY SCALE

Group	Q25	Q50	Q75	P99
1	6.0	10.3	16.8	32.0
2	5.8	10.1	14.7	32.0
3	4.8	8.5	12.6	31.0
4	6.3	8.8	14.2	29.0
5	10.5	15.3	24.3	56.0
6	9.0	14.4	20.9	95.0
7	8.9	18.3	30.3	76.0
Total N, 1098				

*Group 1:* N. 125. This is the original standard population, male, modal age 18, graduating senior boys, highly varied in intelligence, socio-economic status, cultural and racial background.

*Group 2:* N. 300. General college population approximately equal male and female, all ages and varied religions.

*Group 3:* N. 157. Regular church goers, white protestant denominations. Male and female, ages to 65.

*Group 4:* N. 101. Police officers in the Los Angeles area who are attending college to better fit them for their work.

*Group 5:* N. 160. Male inmates in a California State Youth Prison. Modal age 20.

*Group 6:* N. 145. Adult male inmates, Los Angeles County Jail. All ages.

*Group 7:* N. 110. Adult female inmates, Los Angeles County Jail. All ages.

As to additional validity measures, it can be seen in Table 1 that both adult female and male prison groups present approximately identical data



compared to the youth prison groups. The results for the prison groups are highly successful in indicating significantly different distributions than standard populations.

An additional validity measure was obtained by the sample of regular church goers as a high conforming group. They present an homogeneous distribution on the conformity end of the scale. In addition, another conforming group was selected from sworn police officers who are attending colleges in the Los Angeles area for the purpose of becoming more adequate in their work. They also indicate a high conforming distribution.

Tests of significance between groups are not shown; however, the following groups have significant mean differences as measured by both parametric and non-parametric tests of significance. Group 1 with 3, 4, 5, 6, 7, 8; Group 2 with 3, 4, 5, 6, 7, 8; Group 3 with 5, 6, 7, 8; Group 4 with 5, 6, 7, 8.

#### G. SUMMARY

1. Social conformity was defined by 6 determiners: (a) Moral values; (b) positive goals; (c) reality testing; (d) ability to give affection; (e) tension level; and (f) impulsivity. These determiners provided the basis in the content for the construction of two forms of an attitude scale using the direction of perception technique of attitude measurement to measure social conformity.

2. The results of 225 *IIRIs* of high school graduating males were analyzed indicating the operation of a direction of perception. In addition, differences in intelligence, socio-economic status, and cultural background also proved insignificant.

3. A comparison was made of 160 *IIRIs* of male inmates in a youth prison, designated as a non-conformity group, with 125 cases of the standard population. The scoring was based upon 37 weighted items determined by the empirical distribution of responses to these items by the standard population. A significant difference in mean scores, in the desired direction, was found.

4. Further weighting of 17 items out of the 37 was based upon the empirical distribution of responses to these items by the prison group. The 160 prison group *IIRIs* were rescored and compared with a new sample of 108 male and female college students' records scored with the new key. It had been determined that neither sex, age within limits, nor religion would contaminate responses to the scale. This comparison indicated a large significant difference in the desired direction.

5. The non-conformity group was also compared with 101 Police Science students designated as a conformity group through interest and intent. A large significant difference provided additional assurance of the validity of the scale of its assumptions.

6. Additional criterion populations were obtained to demonstrate the continuous ability of the scale to differentiate between conformity and non-conformity groups. These groups were (a) 157 regular church goers, mixed sex, all ages; (b) 110 adult female inmates of the Los Angeles County Jail; (c) 145 adult male inmates of the Los Angeles County Jail.

7. Possible use of this scale is suggested for the following areas: (a) as a research tool; (b) as a demonstration tool of a socio-psychological concept and indirect method of attitude measurement; and (c) for screening and diagnostic purposes where social conformity as defined reflects desirable traits.

#### REFERENCES

1. BERNBERG, R. E. The direction of perception technique of attitude measurement. *Int. J. Opin. & Attit. Res.*, 1951, 5, 397-406.
2. ———. Socio-psychological factors in industrial morale: I. The prediction of specific indicators. *J. Soc. Psychol.*, 1952, 36, 73-81.
3. PETERS, C. C., & VAN VOORHIS, W. R. *Statistical Procedures and Their Mathematical Bases*. New York: McGraw-Hill, 1940.
4. SHERIF, M. *An Outline of Social Psychology*. New York: Harper, 1948.

*Department of Psychology*  
*Los Angeles State College*  
*855 North Vermont Avenue*  
*Los Angeles 27, California*

## A WORK SAMPLE TEST FOR FOREIGN LANGUAGE PROGNOSIS\*

*Language Aptitude Project, Harvard University*

---

STANLEY M. SAPON

---

### A. INTRODUCTION

The notion of the work-sample as a prognostic instrument has high face validity, and has received much attention, both in theory and practice, as well as application in many areas. The value of its extension to the area of foreign language prognosis has been indicated (4), but its application, as evidenced by the work of several investigators (3, 5, 6, 7), has presented some important problems. Efforts in areas of more tangible realization, such as those involving motor performance, have permitted reasonable approximations of the goal task in work-samples which require, from a practical testing standpoint, little specific training for performing in the test situation.

The complex represented by foreign language learning, the diversity of achievement criteria manifested in the wide variety of languages taught in present school systems, as well as the problem of testing naïve individuals in reasonable times, makes the application of work-sample techniques to foreign language prognosis relatively difficult.

One solution, in the form of instruction in an artificial "language" and subsequent measurement of learning, has yielded some success (3, 5, 7). It is the purpose of this paper to outline the theoretical desiderata of this solution, and to describe the design and application of one such "language" on the basis of current linguistic theory and present day tendencies in foreign language instruction. Some preliminary validity data on its use in a prognostic instrument will be presented.

### B. THE PROBLEM

The utilization of a completely novel, constructed "language," rather than one of the extant tongues, is indicated by the need to level any possible differentials of language experience of the student that might tend to distort the value of his scores, as well as to avoid the equivalent danger of

---

\*Received in the Editorial Office on August 13, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

employing a language possessing similarities of structure or vocabulary to the language or languages already known to the student. There is a further factor of value involved in the ability of the test constructor to control the degree of difference from English (considered here as the native language of a monolingual American student, and hence representing one dimension of difficulty), and to grade these degrees of difference throughout the test.

The following are thought by the writer to be primary desiderata in the design of a work-sample test of foreign language learning performance:

1. Verisimilitude of linguistic structure. If the artificial "language" is to serve as a representative work-sample, it must reflect linguistic phenomena actually found in the languages of the world, or more realistically, at least in the languages commonly studied. It should present complexities and resolutions of code structure as reflected in the development of human language, rather than in the construction of an abstract code-problem. Artificial "languages" such as Esperanto, as used in the Symonds Foreign Language Prognosis Test, Form *A* (7), the George Washington University Series Language Aptitude Test (3), and the Iowa Foreign Language Aptitude Test, Form *M* (5), represent only a partially satisfactory solution, inasmuch as "languages" of this type are the result of dedicated efforts to reach a linguistic compromise, with ease of learning for speakers of English a considerable factor in their design. Note could well be made of the complete regularity of adjective-adverb formation, verb conjugations, etc., lack of grammatical gender, and a completely analytic, English-like word order.

2. Minimization of structural similarity to English, and the converse, broad sampling of structural characteristics differing from English. This demand stems from the observation that opportunity for positive transfer with regard to specific language features is more a fortuitous occurrence in actual foreign language study than a regular phenomenon.

The linguistically unsophisticated test constructor runs the risk of unrealistically facilitating the transfer of like structural elements even in the construction of an artificial "language" other than the Esperanto type. Interesting evidence of this can be seen in the "languages" used in the *ACE* Psychological Examination, 1934 edition (1), and the Symonds Foreign Language Prognosis Test, Form *B* (7). Ostensibly and apparently two different, unrelated "languages," analysis reveals that they are identical in structure, with a point for point correspondence in phonemic scheme, morphology and syntax, employing only a different set of speech sounds to



cover their structural skeletons. Changing the sounds in a fixed structural pattern does not create new "languages," any more than substituting terms in an equation creates a new equation, regardless of how different the end products may appear. The inter-relationship of the two "languages" cited above is matched by a marked similarity to English in terms of correspondence of structural pattern.

3. Pedagogical verisimilitude. Wolfe has observed with regard to training programs in general that "... practice materials should vary in as many dimensions and over approximately as wide a range as will the situations to be encountered when the learning is to be applied" (8, p. 1272). In terms of a prognosis problem where one is concerned not only with the aptitudes of the individual, but also with the nature of his response to a training program, the training and measurement techniques applied in the work-sample should approximate, by these standards, those likely to be found in actual course-work. Herein should also be considered the rôle of various factors in language learning. Rote memory, while of obvious significance in some types of foreign language training, should not be depended upon to the exclusion of other factors. The growing acceptance of more direct methods of instruction calls for an equivalent emphasis on inductive learning.

4. Practical considerations of complexity of material and testing time. Although this would appear to be a routine problem of organization of test material, the structure and time-limits of the test depend in large measure on the "language" itself. The Luria-Orleans Modern Language Prognosis Test (6), for example, using short, traditional lessons in Spanish and French as the work-sample material, requires a total of 76 minutes to survey some of the basic structural characteristics of these two languages. The "language" of the work-sample must, within the framework of previously stated needs, be of sufficiently compact structure to permit adequate instruction and practice in a short time period, yet still allow the constructor to manipulate varying degrees of complexity of the material. In addition, the problem of providing some degree of motivational impulse towards optimum performance in the test must also be considered in the test design.

### C. THE LANGUAGE

The language, called Temtem, is briefly described below.

#### 1. Phonology

Five vowel phonemes: /a/, /e/, /i/, /o/, and /u/. Nine consonant phonemes, based primarily on voiced-voiceless oppositions: /m/, /p/, /b/, /k/,

/g/, /t/, /d/, /s/, and /z/. Stress: all free forms have paroxytonic stress. Intonation: intonation is structurally significant in two instances; falling tone before a pause marks termination of phrase, and rising tone before a pause indicates copulative relationship between the elements separated by the pause.

## 2. Morphology

All free forms in zero inflection are bisyllabic, of the *CVCV* type. Bound particles are monosyllabic, of no fixed consonant-vowel pattern. Gender in nouns is natural, and not signified. Natural gender in verb forms is expressed by suffix, /-m/ masculine, /-t/ feminine. Number is represented in terms of "singular," "dual," and an indeterminate "plural," and is indicated by prefix. Nominal prefixes of number are /0-/ , /a-/ , and /az-/ for singular, dual and plural, respectively. Verbal prefixes in the same order are: /0-/ , /i-/ , and /iz-/ . Case: Temtem has two cases, functionally an "active" and an "acted upon," referred to as nominative and oblique. Oblique forms are derived from the nominative by substituting /-u/ for final vowel.

## 3. Syntax

Word order is completely free, except for adjectives, which are uniformly in postposition to the noun. Particles in analytic function precede the morph they affect. Copulative function, as has been indicated, is achieved by intonation. Since the gender and number of the actor are indicated within the body of the verb form, redundancy can frequently be controlled by omission of the actor noun.

## 4. Lexicon

Two verb "roots," two adjectives of color, one conjunctive particle, and seven nouns.

This "language," which was designed in accord with the requirements stated in points one, two and four of the preceding section, has been incorporated into an instrument, described below, to meet the requirements of points three and four.

## D. THE INSTRUMENT

The test is divided into three sections, each including an instruction and a testing period. Total time of administration is 26 minutes. Instruction is accomplished by direct audio-visual method, employing a pre-recorded tape and projected lantern slides. With the exception of the introductory remarks and incidental instructions, no English is used during the test. Students are informed via the tape that they will learn a little about a new

and unfamiliar language called Temtem, "a simple form of the kind of languages spoken by several tribes of American Indians," and that they will have to "listen carefully to differences between words and sentences [as indicated by the pictures on the screen] and figure out for themselves the reasons for these differences."

A picture appears on the screen, and the instructor's voice is heard pronouncing the word or sentence that describes the picture. The picture remains on the screen approximately 15 seconds, during which time the students repeat aloud the instructor's utterance.

Proceeding in this fashion, Lesson I, with 30 slides, presents the nouns and develops sequentially the concepts of number and addition, and the copulative artifact (see Figure 1). Following this, learning is measured in

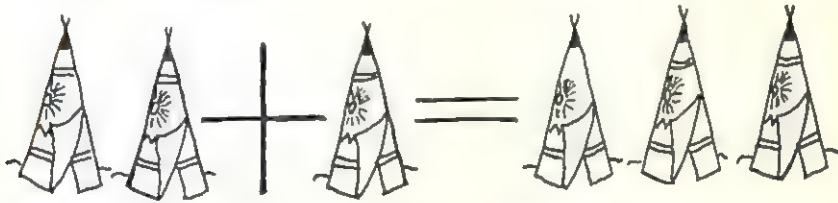


FIGURE 1

Sample of instruction slide. The Temtem utterance that accompanies it is "*atipa si tipa 'aztipa.*"

terms of aural comprehension. For each of 10 numbered items in the printed answer booklet are found four alternative pictures, and the student is required to mark the picture that corresponds to the Temtem utterance heard on the tape playback.

Lesson II repeats the material of Lesson I and adds the colors "red" and "green." Test II, with 10 items, follows the pattern of Test I.

Lesson III repeats the material of Lessons I and II, and introduces the verbs and the consequent use of the oblique case. The material in this lesson is manipulated so that a high degree of sentence complexity is achieved, yielding sentences like "*izkopim adiku si azkimu azpima isami*" (three red men are shooting [with bow and arrow] two horses and many cows). As shown in Figure 2, this complexity is matched by items found in the final test.

## E. VALIDATION

### 1. Sample

On February 8, 1954, this test was administered to a group of 111 airmen at an Air Force base, as one of a larger battery of aptitude tests cur-

rently being developed at the Harvard Language Aptitude Project. This group had volunteered for an intensive four-day course in Mandarin Chinese, successful completion of which would qualify them to take an eight-months' course in Mandarin Chinese at a civilian university. The test

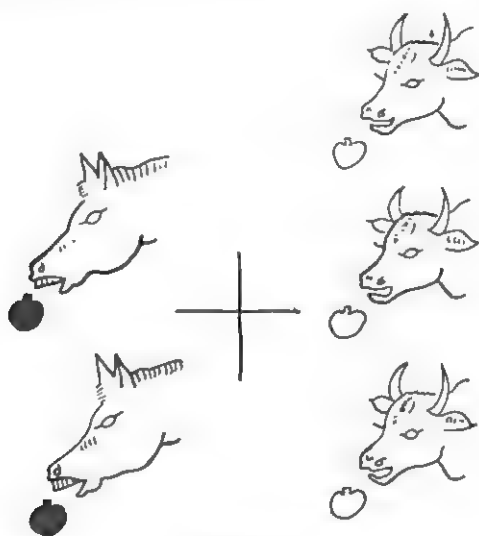


FIGURE 2

Correct answer to Item 36. The test booklet, printed in three colors, shows red and green apples. The Temtem utterance heard is "*igosim adika mibu iduki 'si izgosit mibu isami azkima.*" The three distractors present varying combinations of horses and cows eating red and/or green apples.

was administered in the morning, and following the completion of the administration of the entire test battery, the training began.

## 2. Criterion Bases

Airmen taking the trial course were to be selected on the basis of achievement consistent with the aural-oral emphases of the course. The criteria for success in the course as described by Frith (2) are:

1. Retention of 90 per cent of the vocabulary of the trial course at the end of the fourth day, as measured by tests administered orally with written translation responses.
2. Comprehension of 90 per cent of the utterances using the vocabulary and structural patterns introduced during the trial course, as measured by tests administered orally with written translation responses.
3. Ability to *recognize* the phonological distinctions of Mandarin, as



measured by dictation tests in which the utterances were recorded in romanized script with appropriate tone designators.

4. Ability to *make* phonological distinctions of Mandarin, as measured by instructor's ratings.

5. Ability to recombine the vocabulary within the limits of the structural patterns of the trial course, as measured by instructor's ratings.

On this basis, 33 students were selected, 30 of whom went on to continue their studies at the civilian university. The criterion scores made available to the project were average percentage grades for each student.<sup>1</sup> Information was also included as to which students had been selected for the eight-months' course, which students failed the course, and which students withdrew voluntarily.

#### F. RESULTS

Validity of the instrument as a predictor was determined by using the total score on the test as the predictor variable, and a product-moment correlation with the normalized academic grade as the criterion yielded  $r = .53$ , significant at the one per cent level of confidence. Excluding cases of voluntary withdrawal,  $N = 80$ . The reliability of the test was .68, as determined by the Kuder-Richardson formula 20.

Although the obtained reliability may be considered low, it appears to be due to the small number of items in the test, and to the error contributed by several ambiguous items, as revealed by an item analysis. In an attempt to raise its reliability, the instrument is being modified by increasing the number of items in the test, patterned after those items which best discriminated between good and poor students, and eliminating those items found to be ambiguous.

#### G. SUMMARY AND CONCLUSIONS

The notion of a work-sample test as a predictor of performance has been considered useful by many investigators. The application of this notion to foreign language prognosis, however, raises special problems, to which the artificial "language" has been offered as a solution. The optimum requirements for such a test are discussed from the standpoint of linguistics, and pedagogical and test-construction considerations, and an instrument employ-

<sup>1</sup>Since the distribution of the grades was considerably skewed, they entered into the computations both as "raw" grades and as "normalized" grades. The normalized grade was the normal deviate, expressed in the form of a *T*-score, corresponding to the mid-percentiles of class intervals of the frequency distribution.

ing an artificial "language" designed on the basis of the aforementioned criteria is described.

The results of its application to a student sample undergoing a trial course in Mandarin Chinese seem to confirm the predictive value of this type of foreign language work-sample, and suggest its further development.

#### REFERENCES

1. American Council on Education Psychological Examination. 1934 Edition, prepared by Thurstone, L. L., & Thurstone, T. G. Washington, D. C.: American Council on Education, 1934.
2. FRITH, J. R. Selection for language training by trial course. *Report of the Fourth Annual Round Table Meeting on Linguistics and Language Teaching*, No. 4. Washington, D. C.: Georgetown Univ. Press, September, 1953.
3. George Washington University Series Language Aptitude Test. Prepared by Hunt, Thelma, and others. Washington, D. C.: Center for Psychological Service, 1929.
4. HENMON, V. A. C., *et al.* *Prognosis Tests in the Modern Foreign Languages*. Publication of the American and Canadian Committee on Modern Languages, Vol. 14. New York: Macmillan, 1929.
5. Iowa Placement Examinations, Series FA-2, Form M, Foreign Language Aptitude. Constructed by Stoddard, G. D., and revised by Cochran, G., Nielson, J. R., and Stuit, D. B. Iowa City: Bureau of Educational Research and Service, State University of Iowa, 1941.
6. LURIA, M. A., & ORLEANS, J. S. Luria-Orleans Modern Language Prognosis Test. Yonkers-on-Hudson, New York: World Book, 1930.
7. SYMONDS, P. M. Foreign Language Prognosis Test. New York: Bureau of Publications, Teachers College, Columbia University, 1930.
8. WOLFE, D. Training. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 1267-1286.)

*Language Aptitude Project*

*McKay Laboratory, Harvard University*

*9 Oxford Street*

*Cambridge 38, Massachusetts*

## DIFFERENTIAL PROBLEM SEQUENCES AND THE FORMATION OF LEARNING SETS\*<sup>1</sup>

*Department of Psychology, Emory University*

C. L. DARBY AND A. J. RIOPELLE

### A. INTRODUCTION

The significance of many factors for the formation of learning sets (1) remains to be determined. One factor not yet evaluated is the sequence with which problems are presented to the experimental subjects. Problem sequences involving interspersed trials of several problems have been found effective for evaluating the intellectual capacity of chimpanzees (2) and of normal and brain-operated monkeys (3), and it appears reasonable that the nature of the problem sequence might affect the rate with which learning sets are formed.

In the present experiment interest is directed toward a comparison of the effects of the successive and the concurrent problem sequences. Under the successive procedure each problem is presented for six consecutive trials, whereas in the concurrent procedure one trial is given on each of several problems before the second trial of any problem is presented.

### B. METHOD

#### 1. *Subjects*

Seven adolescent and young adult rhesus monkeys served in this experiment. All had undergone taming procedures and all had learned 10 preliminary object-quality discrimination problems.

#### 2. *Apparatus and Test-Trial Procedures*

The Emory University adaptation of the Wisconsin General Test Apparatus (1) was used. This device and detailed test trial procedures have previously been described (4).

Two stimulus objects comprise a problem and are placed simultaneously on a sliding tray confronting the monkey. Both objects cover food wells,

\*Received in the Editorial Office on August 13, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This research was supported in part by a research grant (M-589) from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

one of which is baited. When the animal displaces the correct object he obtains food. If he displaces the other object, however, the tray is withdrawn and reset for the next trial. The non-correction procedure was used throughout.

### 3. *Experimental Design*

The experiment was conducted in two phases. During Phase I all animals were administered a total of 240 problems, six trials per problem, 10 problems per day. In this first phase the animals were divided into two groups. Three animals [4, 8, 21] were assigned to the Successive Problem Group. For this group the problems were presented consecutively; each new problem beginning only after all six trials of the previous problem had been completed. The remaining four animals [1, 3, 7, 18] were assigned to the Concurrent Problem Group. These animals were given Trial 1 for all 10 problems before beginning Trial 2. Trial 2, in turn, was presented for all 10 problems before beginning Trial 3. The successive presentations of the stimuli for a single problem occurred in an unsystematic sequence.

Phase II began immediately after Phase I and was pursued for a total of 100 problems. During this phase all animals were tested under both the successive and the concurrent sequences. The two conditions were alternated on successive days.

## C. RESULTS

### 1. *Phase I*

Figure 1 shows the results of Phase I. The two graphs of this figure portray performance on Trial 2 and on Trial 6 for the Successive and the Concurrent Groups. The points on the graphs represent per cent correct choices in successive blocks of 40 problems.

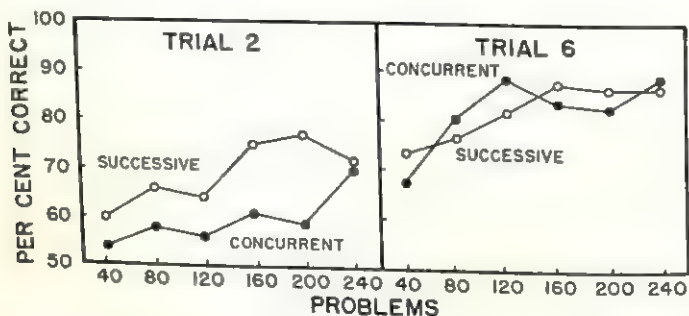


FIGURE 1

PER CENT CORRECT RESPONSES ON TRIAL 2 AND ON TRIAL 6 FOR THE SUCCESSIVE PROBLEM GROUP AND THE CONCURRENT PROBLEM GROUP



Three things can be noted from these graphs. First, both curves of each panel show an upward slope, thereby indicating that both groups progressively learned the successive problems more rapidly. Secondly, performance on Trial 6 is superior to that of Trial 2 for both groups at comparable stages of training. Thirdly, the performance of the Successive Group exceeded that of the Concurrent Group, although this difference fell short of the usually required levels of statistical significance.

The data of Phase II, in which each animal was tested under both problem sequences, were pooled and analyzed by trials to elucidate the principal differences. Such data are shown in Figure 2 which presents the

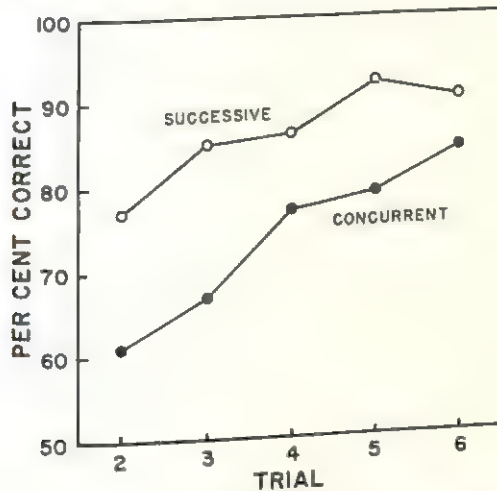


FIGURE 2  
PER CENT CORRECT RESPONSES FOR ALL ANIMALS ON TRIALS 2-6 OF THE PROBLEMS  
ADMINISTERED ACCORDING TO THE SUCCESSIVE PROCEDURE AND THE  
CONCURRENT PROCEDURE DURING PHASE II

percentage of correct choices on Trials 2-6. Here again, performance under the successive condition is superior to that under the concurrent procedure. Also, the curve for the concurrent procedure begins lower and rises more rapidly within the six trials than does the other curve. This finding agrees with the results of Phase I in that differences between the two conditions were greatest for Trial 2.

Statistical treatment of the results of Phase II yielded several significant variance estimates which support the interpretations placed on the data. First, the difference between procedures was highly significant (.01 level) even when tested against the largest interaction term involving procedures. In addition, differences among animals and among trials were also significant.

Finally, the interpretation that the slopes of the two curves in Figure 2 are different is supported by the finding of a significant interaction between trials and procedures.

#### D. DISCUSSION

The principal result of this investigation was the inferiority of the concurrent procedure when compared with the successive problem sequence. This finding is in agreement with earlier data by Hayes, Thompson, and Hayes (2) on chimpanzees which showed that performance deteriorates if the subjects are presented too many problems to handle concurrently. It is also in line with expectations based upon proactive and retroactive inhibition since the concurrent procedure maximizes interproblem interpolation and interference.

#### E. SUMMARY

This experiment compared the discrimination learning performance of seven rhesus monkeys tested under a successive problem sequence and under a concurrent problem sequence. The successive problem sequence resulted in superior discrimination learning performance.

#### REFERENCES

1. HARLOW, H. F. The formation of learning sets. *Psychol. Rev.*, 1949, **56**, 51-65.
2. HAYES, K. J., THOMPSON, R., & HAYES, C. Discrimination learning set in chimpanzees. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 99-104.
3. RIOPELLE, A. J., HARLOW, H. F., SETTLAGE, P. H., & ADES, H. W. Performance of normal and operated monkeys on visual learning tests. *J. Comp. & Physiol. Psychol.*, 1951, **44**, 283-289.
4. RIOPELLE, A. J. Facilities of the Emory University Primate Behavior Laboratory. *J. of Psychol.*, 1954, **38**, 331-338.

*Department of Psychology*  
*Emory University*  
*Emory University, Georgia*

## THE PERFORMANCE OF ADULTS UNDER DISTRACTION STRESS: A DEVELOPMENTAL APPROACH\* <sup>1, 2</sup>

*Worcester State Hospital and Clark University*

---

STANLEY H. LOFCHIE<sup>3, 4</sup>

---

### A. INTRODUCTION

The advent of World War II served to make problems of stress a central issue in psychology. In addition to such official undertakings as the Aviation Psychology Program (12) and the O.S.S. "Assessment of Men" (13) there have been numerous other investigations in this area by experimental, clinical, and applied psychologists. The findings reported in these studies, however, are notable for their lack of agreement. Lazarus, Deese, and Osler (10) point out that much of the inconsistency of these findings may be attributed to a lack of adequate analysis of *S*'s task. Consequently, the present study, which attempts to ascertain what factors underlie performance under stress, employs an experimental technique designed to facilitate an analysis of *S*'s task. The technique utilized is a variation of one of the most widely used and one of the least complex reported in the literature: a simple psychomotor task with distraction induced by strong extraneous stimuli.

The Rorschach technique was used as a means of measuring certain psychological factors presumed to be associated with performance under such conditions.

### B. STATEMENT OF THE PROBLEM

The purpose of this study is an investigation of those factors which underlie performance under stress. Towards this end, performance under distraction stress was analyzed in terms of genetic theory as outlined by Werner (19). The basic principle of this theory is that behavior can be viewed as a developmental process which proceeds from an undifferentiated stage to one of discreteness and then to a developmentally mature level characterized

---

\*Received in the Editorial Office on August 17, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was supported by a Research Grant from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

<sup>2</sup>Based on a dissertation accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Clark University, 1953.

<sup>3</sup>The writer wishes to express his thanks to Dr. Leslie Phillips and the staff of the Worcester State Hospital Research Service for their advice and criticism.

<sup>4</sup>I wish to express my appreciation to John E. Lane.

by hierarchic integration, i.e., an appropriate organizational pattern. With reference to performance under distraction stress genetic theory implies that developmentally mature individuals (i.e., those with a greater ability to appropriately organize the behavioral field) will tend to respond selectively to that which is relevant.

In the stress situation used in the present study, irrelevant (distracting) stimuli are presented simultaneously with the relevant (task) stimuli. *S* must discriminate between these two sets of competing stimuli and he must also achieve a functional separation between them in order that the non-essential does not impair his reaction to the essential. He must organize a perceptual-motor stimulus pattern in accordance with the requirements of given task goals despite the presence of stimuli that conflict with these requirements.

To test the implications of genetic theory for performance under distraction stress it is clear that a measure of developmental maturity is necessary. Developmental maturity, as reflected in the appropriate organization of the perceptual field, has been studied within the genetic framework chiefly by means of the Rorschach test (6, 7, 8, 14, 17). This work has demonstrated that developmental changes in perceptual organization occur both with increasing age and degree of psychopathology. In terms of the studies cited appropriately organized Rorschach responses are *F+* and require specific form. Further, they are not based primarily on the outline of a solid blot area; rather, they involve either the analysis and appropriate reintegration of a solid blot area or the appropriate integration of blot areas that are spatially distinct.

An analysis of the processes involved in the development of these responses in the Rorschach test suggests that they are analogous to the processes involved in adequate performance under distraction stress.

There are various features of the Rorschach blot which make difficult the production of formally adequate integrated responses. For example, integrated responses require a discrimination between those blot elements that are relevant and those that are not relevant to any given content. At the same time, however, the organizing process may be obstructed by minor discrepancies between the blot and the percept. The two principal factors that give rise to such discrepancies are form and color. Thus not infrequently a subject may respond to the effect that a given area "could be a man if only the toes weren't shaped like that"; or "that would make a perfect bear—but it can't be—it's red." In a very real sense then, these discrepancies between blot and percept may be regarded as elements which distract the subject and obstruct the operation of an integrative activity.



Siipola (18) specifically considers the influence of color on the Rorschach in terms of distraction—pointing out that the colored cards are difficult to organize for adults not because of the presence of color per se but rather because a discrepancy exists between a given color and the percept implied by the form. The frequent coexistence on the Rorschach of form discrepant with color demonstrates that on this test relevant and irrelevant stimuli are interwoven. The task of the subject is to organize only the relevant stimuli into an adequate response.

There are, of course, different behavioral requirements in the Rorschach and the performance task. The pertinent identity between the two tasks is the integrating perceptual function that is requisite for high level performance in both.

The present study, consequently, tests the hypothesis that the motor performance of individuals who evidence greater ability to appropriately organize the perceptual field (as measured by the Rorschach test) will be less affected by distraction stress than will the motor performance of individuals of lesser ability to appropriately organize the perceptual field.

### C. PROCEDURES

#### 1. *Subjects*

The experimental population consisted of 35 white, American-born, normal, adult male industrial workers between the ages of 34 and 45. All Ss were volunteers and were paid for their services.

#### 2. *Order of Experimentation*

Each S spent two consecutive mornings in experimentation. In no case were the two techniques (Rorschach and motor task) administered on the same day. Seventeen of the 35 Ss were administered the Rorschach first, and for the other 18, it was administered on the second day of experimentation.

The Vocabulary sub-test of the 1916 Stanford-Binet immediately preceded all Rorschach administrations. It was used to control for the possible influence of intelligence level on the S's performance with the experimental procedures.

#### 3. *The Motor Task*

A modified Dunlap steadiness apparatus (3) was utilized in the motor task. In the present apparatus, a stylus was fixed so that it could not be withdrawn from a hole. At the same time the stylus was delicately balanced so that S's constant attention was required to prevent it from touching the sides of the hole. In the present study only one hole size (one-eighth

of an inch in diameter) was used. *S* always knew when the stylus touched the sides of the hole because such contact extinguished a small light which illuminated the hole. An electric timer provided a measure of error-free time for each trial, i.e., a measure of that portion of the trial during which the stylus was not in contact with the sides of the hole. The amount of error-free time provided the basis for the evaluation of performance under distraction stress.

*E* introduced the task as an "eye-hand coordination test," and the distracting stimuli were not specifically mentioned. However, it was emphasized that *S* was to continue at his task "no matter what may be happening around you." This was done to obviate the possibility of a total interruption of performance with the introduction of the distracting stimuli.

Each *S* was given a demonstration of the apparatus and was accorded practice sufficient to allow him to fully grasp the nature of the task. In all but a few cases one eight-second trial proved adequate for this purpose. Then, in the experiment proper each *S* was given a total of 30 trials, each of eight-second duration. There was a half-minute rest period between trials. Extraneous stimuli (air blast directed at back of *S*'s neck, bright flashing overhead light, loud horn and bell) were introduced during 10 of the last 20 trials. These trials were designated as distraction trials and occurred according to a pre-arranged irregular pattern that was held constant for all *S*s. Trials one through 10 were designated as pre-distraction trials and the remaining 10 trials were called distraction-free trials.

It has been noted that prestress level of performance is one of the determinants of subsequent performance (10). Accordingly a Pearson  $r$  was calculated between the group's performances under pre-distraction and distraction conditions. This correlation co-efficient was  $+.87$ , significant beyond the .01 level of confidence. A regression equation based on this correlation was then used to predict each *S*'s performance on the distraction test. (Inspection of the data indicated that the assumption of equality of variance along the arrays was satisfied.) Those *S*s whose actual performance exceeded this predicted score were then designated as "good performers," those whose performance failed to meet expectancy were classified as "poor performers," and those whose performance on the distraction test matched the predicted score were called "median performers."

#### D. THE RORSCHACH TEST

All Rorschach protocols were obtained by the individual administration of the standard series of 10 cards. Administration was essentially in accordance with the principles of Beck (1). The protocols were scored in terms of a

developmental scoring system devised by Friedman (7) and empirically validated in other studies (6, 8, 14, 17). In the present study the extent to which appropriately organized Rorschach responses prevailed in a given Rorschach protocol was taken as a measure of organizing capacity, and, hence, perceptual maturity.

A full description of the scoring system can be found elsewhere (16). For the purposes of the present paper it is sufficient to note that scores in this system denoting organizing capacity refer to Whole and Usual Detail responses. In the present study, therefore, the Index of Perceptual Maturity was calculated as follows:

$$\frac{\text{Organized W} + \text{Organized D}}{\text{W} + \text{D}} \times 100.$$

The reliability of the application of this new development scoring system in the present study was checked by an independent scoring of a sample of 10 records taken at random from the protocols involved in the present study.<sup>5</sup> A total of 197 responses were involved in the 10 records. The obtained percentage of agreement was 94.4. This compares favorably with the mean agreement (93.05 per cent) of the six pairs of scorers involved in the first study using this scoring system (7). None of the disagreements involved any of the scores that contribute to the Index of Maturity.

To provide a further check on the reliability of the Index the present writer computed this score for each of the 30 normals involved in Friedman's study (7). The Maturity Index scores of the present group of Ss were then compared with the scores of Friedman's group by means of a chi-square analysis. No significant difference was found.

The experimental hypothesis was that this Index of Perceptual Maturity would be positively related to performance under distraction stress.

### E. RESULTS

The difference in organizing ability between good and poor performers was tested by calculating a median for the combined Maturity Index distribution and comparing the two groups for the number of cases falling above and below this point. All median cases were excluded. Fisher's (5) exact method was used to evaluate the resultant four-fold contingency tables.

The hypothesis to be tested was that the Index of Perceptual Maturity would be positively related to performance under distraction stress. This hypothesis was strikingly confirmed ( $p < .001$ ). The contingency table is presented in Table 1.

<sup>5</sup>I wish to express my appreciation to Robert C. Misch, who independently scored these Rorschach records.

TABLE 1  
THE RELATIONSHIP BETWEEN PERCEPTUAL MATURITY AND PERFORMANCE UNDER STRESS  
( $N = 30$ )

Group	Maturity index	
	Below median	Above median
Good performers	2	13
Poor performers	13	2

$P = < .001.$

The data were examined for the influence of variables which were not directly pertinent, i.e., such variables as intelligence, education, occupation, number of Rorschach responses, emphasis in Rorschach location choice, and experimental schedule. The relationships between each of these factors and the Index of Perceptual Maturity and between each of these factors and performance scores were tested. The results were as follows: (a) The Index of Perceptual Maturity was not significantly related to any of these variables and showed only a tendency ( $p = .19$ ) to be related (positively) to intelligence. (b) Performance scores were not significantly related to any of these variables—nor were any marked trends discernible. The  $P$ -value for the (positive) relationship between intelligence and performance was .45.

The relationship between the Index of Perceptual Maturity and performance scores was then calculated with controls instituted for level of intelligence. Cronbach's (2) technique of roving medians was used so that the Rorschach scores were compared with performance scores for groups in which the range of intelligence was restricted. The population was divided into three groups of approximately equal numbers. These represented the upper-middle and lower levels of intelligence found in the experimental population. When the effects of intelligence were controlled in this manner, the obtained  $P$ -value for the relationship between the Index of Perceptual Maturity and performance scores was .01.

Consequently, it seemed that the obtained relationship between perceptual maturity and performance under distraction stress was dependent neither on the factor of intelligence nor on the other variables—i.e., response production, location choice, education, occupation, and schedule.

## F. DISCUSSION

The results of the present experiment suggest that the common psychological process which underlies performance under distraction stress and formally adequate integrated responses on the Rorschach technique is an ability to restrict responses to only those stimuli in the perceptual field which are relevant.



This psychological process would appear to be important in other experimental situations. In the orientation tests used in the Brooklyn College Perception-Personality Project, for example, *S*'s task was to orient himself or an object in space while being subjected to sensory experience which ran counter to the demands of the task. One typical procedure consisted of presenting in a darkened room a luminous rod and frame both tilted to an angle of 30 degrees. The task of the *S* was to adjust the rod to the true vertical while the frame remained stationary. In terms of the analysis of the present experimental conditions the frame in the orientation task can be considered as irrelevant or distracting stimuli while the rod constitutes the relevant stimuli. *S* is required to achieve a functional separation between the two. In fact Witkin *et al.* (21) do report that success in this task is dependent on the ability of "keeping-item-separate-from-field." Witkin (20) however has further described these orientation situations as characterized by a conflict in which "one's own impulses provide one standard for action and the demands and restrictions of the environment another." This observation, moreover, was substantiated by later investigations (21) which demonstrated that ability to perform adequately in these orientation tasks was related to a general capacity for coping with environmental pressures. The evaluation of "coping capacity" was made on the basis of *S*'s behavior and attitudes in school, home, and military service and included such diverse topics as leadership, attitudes towards career, and attitudes toward marriage and sex.

Similarly, it might be supposed that performance under conditions of distraction stress also may be related to an ability to handle environmental pressures. This finds support in a study by Feffer and Phillips (4) in which scoring on a measure of social effectiveness was found to be related to performance under both distraction and failure stress conditions.

### G. SUMMARY

An analysis of the conditions of *S*'s task in a psychomotor performance under distraction stress led to the hypothesis, derived from developmental theory, that those *S*'s scoring higher on a Rorschach Index of Perceptual Maturity would be better performers on a psychomotor task, administered under conditions of distraction stress, than would those *S*'s scoring lower on the same index. Data obtained with a group of normal adult males confirmed this hypothesis.

The implications of this study for the possible relationships between performance under stress and level of social adequacy were discussed.

## REFERENCES

1. BECK, S. J. *Rorschach's Test: Vol. 1. Basic Processes.* (Sec. Ed. Rev.) New York: Grune & Stratton, 1949.
2. CRONBACH, L. J. Statistical methods applied to Rorschach scores: a review. *Psychol. Bull.*, 1949, **46**, 393-429.
3. DUNLAP, K. Improved forms of steadiness tester and tapping plate. *J. Exp. Psychol.*, 1921, **4**, 430-433.
4. FEFFER, M., & PHILLIPS, L. Social attainment and performance under stress. *J. Pers.*, 1953, **22**, 284-297.
5. FISHER, R. A. *Statistical Methods for Research Workers.* (Eleventh Ed. Rev.) New York: Hafner, 1950.
6. FRANK, I. H. A genetic evaluation of perceptual structurization in certain psychoneurotic disorders by means of the Rorschach technique. Ph.D. dissertation, Boston Univ., Boston, Mass., 1952.
7. FRIEDMAN, H. Perceptual regression in schizophrenia: An analysis by means of the Rorschach test. Ph.D. dissertation, Clark University, Worcester, Mass., 1950.
8. HEMMENDINGER, L. A genetic study of structural aspects of perception as reflected in Rorschach responses. Ph.D. dissertation, Clark University, Worcester, Mass., 1951.
9. LANE, J. E. The effect of a stressful situation on the psychomotor learning of schizophrenic and normal subjects. M.A. Thesis. Clark University, Worcester, Mass., 1951.
10. LAZARUS, R. S., DEESE, J., & OSLER, S. F. The effects of psychological stress upon performance. *Psychol. Bull.*, 1952, **49**, 293-317.
11. LOFCHIE, S. H. The performance of adults under distraction stress: a developmental approach. Ph.D. dissertation, Clark University, Worcester, Mass., 1953.
12. MELTON, A. W. Apparatus tests. *AAF Aviat. Psychol. Prog. Res. Rep.*, 1947, No. 4. Washington: U. S. Government Printing Office.
13. OSS ASSESSMENT STAFF. *Assessment of men.* New York: Rinehart, 1948.
14. PENA, C. A genetic evaluation of perceptual structurization in cerebral pathology: An investigation by means of the Rorschach test. Ph.D. dissertation, Boston University, Boston, Mass., 1951.
15. PHILLIPS, L. Personality factors and prognosis in schizophrenia. Ph.D. dissertation, Univ. Chicago, 1949.
16. PHILLIPS, L., & SMITH, J. *Rorschach Interpretation: Advanced Technique.* New York: Grune & Stratton, 1953.
17. SIEGEL, E. L. Genetic parallels of perceptual structurization in paranoid schizophrenia: an analysis by means of the Rorschach technique. Ph.D. dissertation, Clark University, Worcester, Mass., 1951.
18. SIPOLA, E. M. The influence of color on reaction to ink blots. *J. Pers.*, 1950, **18**, 358-382.
19. WERNER, H. *Comparative Psychology of Human Development.* (Rev. Ed.) Chicago: Follett, 1948.
20. WITKIN, H. A. The nature and importance of individual differences in perception. *J. Pers.*, 1949, **18**, 145-170.
21. WITKIN, H. A., HERTZMAN, M., LEWIS, H. B., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. *Personality Through Perception.* New York: Harper (in press).

86 School Street  
Lexington, Massachusetts

## MEASURES OF INTELLIGENCE AND CREATIVITY\*<sup>1</sup>

*Department of Psychology, University of Chicago*

---

BERNARD MEER AND MORRIS I. STEIN

---

### A. THE PROBLEM

This report is part of a larger study that is devoted to an investigation of the psychological and environmental factors related to creativity in the sciences. This paper concerns itself with one aspect of this study—the relationship between intelligence and creativity.

In a previous paper (2) the theoretical framework underlying the project was presented. Briefly, it may be said that creativity is regarded as a *process* which involves the acquisition of fundamental tools or basic knowledge in a field, the developing of hypotheses, the testing of hypotheses, and the communication of results to others. It is apparent that intellectual factors should be among those which affect an individual's creativity.

The question posed for this paper is: What is the relationship between intelligence test scores and creativity for a population of subjects actively engaged in research?

### B. THE CRITERION OF CREATIVITY

It is generally agreed that the creative work is a novel work. Stein has pointed out, however, that in addition to its being novel it must be accepted by a group at some point in time (2, p. 311). This criterion, based on external evaluation, cannot be avoided since there are no objective standards for measuring creativity. In the final analysis, whether one deals with ideas, products, or patents, some *person* or *persons* must decide which ideas or products are the more creative ones.

### C. SELECTION OF SUBJECTS

The subjects in the present study were research chemists employed in the two major departments of the research division of a large industrial organization. Lists of all research personnel were obtained and from these

---

\*Received in the Editorial Office on August 18, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This study was supported in part by a grant-in-aid from the Research Division of Armour and Co., Chicago, Illinois, and in part by a research grant (PHS M-726) from the National Institute of Mental Health, of the National Institutes of Health, Public Health Service.

men were selected who met two criteria: (a) ability to do independent research as judged on the basis of training and/or experience; and (b) being sufficiently well known to the department head so that he could judge their creativity. Females and all men over 50 years of age were omitted to maintain homogeneity of population. Twenty-six subjects in Department *A* and 38 subjects in Department *B* met these criteria.

Individual meetings were held with each of the department heads to discuss the meaning of creativity. The criteria for creativity were somewhat different for the two men. While they both agreed that the quality and quantity of ideas were most important, Director *B* included the ability to do something about one's ideas. Following these meetings, each director ranked his men for creativity. One month later, the department heads repeated their rankings. The reliabilities of these rankings in terms of rho were  $+ .95$  for Department *A* and  $+ .89$  for Department *B*, both significant at the .001 level of confidence. A subject's final creativity score consisted of an average of his two rankings.

A limitation of the criterion was its dependence on the judgment of only one person, the department head.<sup>2</sup> As a check on his objectivity, his rankings were compared with the ratings on creativity obtained from all technical personnel in the research division, one year prior to the time that the department heads made their rankings. At that time, each man rated his superiors, colleagues, assistants, and himself for creativity.<sup>3</sup>

In carrying out this check, the mean ratings obtained from each of these separate sources (superiors, colleagues, assistants) for the subjects ranked in the upper third of the department heads' distribution and those obtained for the men in the lowest third of the distribution were compared. In all but one instance (the ratings obtained from assistants in Department *A*, and here the trend was in the right direction), those men who were ranked high by the department directors had been rated significantly higher (at the .01 level) by the other groups than those men whom the department directors had ranked low. Furthermore, the ratings of the department directors themselves which had been obtained one year prior to their rankings were also significantly higher for their top men than for their low men.

#### D. THE MEASURES OF INTELLIGENCE

The Wechsler-Bellevue Intelligence Scale (*W-B*) and the Miller Analogies Test (*MAT*) were used to evaluate the subjects' intelligence. The

<sup>2</sup>This limitation was inherent in the situation since the department head was the only person who could rate *all* the research workers in his department.

<sup>3</sup>These ratings were not used as the main criterion because all the subjects were not rated by the same persons.



former was selected because of its use with adults and because one could obtain, in addition to an overall *IQ*, measures of specific functions from the subtests (2). The *MAT* was selected because of its power to discriminate subjects at the graduate school level (1).

### E. RESULTS

Before proceeding to the presentation of results regarding the relationship between intelligence test scores and creativity, it is necessary to rule out several other factors—age, amount of time in the present organization, total professional experience, and education—which might confound the data. As indicated in Table 1 the first three factors are not correlated with creativity. A breakdown of educational status (Table 2), however, reveals that significantly more Ph.D.s than non-Ph.D.s are to be found in the upper half of the rankings on creativity. In order to control for possible halo effects due to the Ph.D. degree, the data for Ph.D.s and non-Ph.D.s were analyzed separately. These analyses were carried out in addition to those that were related to the study of the original ranks on creativity (in-

TABLE 1  
THE MEANS FOR AGE, TIME IN PRESENT ORGANIZATION, TOTAL PROFESSIONAL EXPERIENCE, AND THE CORRELATIONS BETWEEN THESE FACTORS AND CREATIVITY

	N	Age		Time in present organization (in years)		Total experience (in years)	
		Mean	r	Mean	r	Mean	r
Ph.D.s							
Dept. A	9	31.8	-.19	3.9	.18	7.1	-.07
Dept. B	16	32.9	.00	2.5	.34	6.9	.20
Non-Ph.D.s							
Dept. A	17	36.6	.30	8.5	-.05	11.7	.27
Dept. B	22	32.6	.05	4.5	-.15	7.7	-.01

TABLE 2  
THE RELATIONSHIP BETWEEN CREATIVITY AND EDUCATIONAL STATUS

		Creativity rank		P
		Upper half	Lower half	
Dept. A	Ph.D.	7	2	<.10
	Non-Ph.D.	6	11	
Dept. B	Ph.D.	12	4	<.05
	Non-Ph.D.	7	15	
Combined A & B	Ph.D.	19	6	<.01
	Non-Ph.D.	13	26	

dependent of educational status). This analysis is referred to below as the "consolidated analysis."

Table 3 indicates the relationship between the measures of intelligence and the creativity rankings.<sup>4</sup> This relationship is clearly significant for the consolidated analysis. However, when education is held constant this relationship does *not* hold for Ph.D.s. Even though the Full and Verbal Scale *IQ*'s of the *W-B* show a trend, this is reversed by the negative correlation obtained with the *MAT*. The latter is probably the more sensitive test since it gives a much greater range of scores with the present population than does the *W-B*. For non-Ph.D.s, the relationship between intelligence and creativity as measured by the Full Scale and Performance Scale *W-B* and the *MAT* is positive and significant.

TABLE 3  
CORRELATIONS BETWEEN VARIOUS MEASURES OF INTELLIGENCE AND CREATIVITY

	<i>N</i>	Full scale	<i>W-B</i> Verbal scale	Performance scale	<i>MAT</i>
Consolidated					
Dept. <i>A</i>	26	.38	.42*	.18	.48*
Dept. <i>B</i>	37	.51**	.38*	.37*	.24
Combined <i>A &amp; B</i>	63	.46**	.39**	.31*	.32*
Ph.D.s					
Dept. <i>A</i>	9	.23	.06	.26	.10 <sup>b</sup>
Dept. <i>B</i> <sup>a</sup>	15	.35	.44	.15	— .29 <sup>b</sup>
Combined <i>A &amp; B</i>	24	.32	.32	.17	— .18
Non-Ph.D.s					
Dept. <i>A</i>	17	.30	.46	.05	.37
Dept. <i>B</i>	22	.48*	.20	.54**	.56**
Combined <i>A &amp; B</i>	39	.40*	.30	.32*	.46**

\*Significant at the .05 level.

\*\*Significant at the .01 level.

<sup>a</sup>The scores of one Ph.D. were omitted because of a language handicap.

<sup>b</sup>Two Ph.D.s, one in each department, left the company before the *MAT* was administered. Thus the *N*'s for Ph.D.s here were 8, 14, and 22 respectively.

Table 4 shows the level of significance of the difference between the mean *IQ*'s of all Ph.D.s and non-Ph.D.s on the *W-B* and *MAT* regardless of their ranking on the creativity variable. The results here indicate that the performance *IQ*'s do not differentiate between the two groups, and while

<sup>4</sup>We are indebted to Dr. Henry E. Daniels, Visiting Professor from Cambridge University, England, to the Department of Statistics at the University of Chicago, for his invaluable statistical aid. He developed a rationale for utilizing ranks as scores, which permitted the use of Pearsonian *r*'s in place of rank-order correlations. (Personal communication.)

the same result is obtained with the *W-B* Full Scale, there is a strong trend in favor of the Ph.D.s. On the more verbal measures of intelligence, the Ph.D.s are clearly superior to the non-Ph.D.s (the *W-B* Verbal Scale is significant at the .05 level while the *MAT*<sup>5</sup> is just short of significance at the .05 level).

TABLE 4  
THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN THE MEAN INTELLIGENCE SCORES FOR  
PH.D.S AND NON-PH.D.S

	Dept. A & B	N	Mean	SD	CR	P
<i>W-B</i>						
Full Scale	Ph.D.s	24	132.25	5.88	1.46	N.S.
	Non-Ph.D.s	39	130.09	5.39		
<i>W-B</i>						
Verbal Scale	Ph.D.s	24	131.88	5.79	2.31	<.05
	Non-Ph.D.s	39	128.23	6.54		
<i>W-B</i>						
Performance Scale	Ph.D.s	24	128.17	8.38	.55	N.S.
	Non-Ph.D.s	39	127.04	7.32		
<i>MAT</i>	Ph.D.s	22	65.68	16.05	1.96	N.S.*
	Non-Ph.D.s	39	57.19	16.24		

\*Just short of significance at the .05 level.

Thus far only the general measures of intelligence have been presented. Table 5 shows the relationship between creativity and more specific types of intellectual functioning as measured by the *W-B* subtests. The combined consolidated analysis resulted in five subtests achieving a significant relationship with creativity. Three of the subtests (Information, Digit Span, and Vocabulary) are part of the Verbal Scale and two of the subtests (Picture Arrangement and Block Design) are part of the Performance Scale. The critical question is whether these subtests continue to differentiate the more or less creative subjects when extent of education is controlled. The combined analysis for Ph.D.s does not yield any significant correlations even though Vocabulary is significant for Department A and Arithmetic for Department B. These results follow the more general findings (Table 3), which showed no relationship between the three *W-B* Scales and *MAT*, and creativity for Ph.D.s. For non-Ph.D.s, two subtests show consistent relationship with creativity: Information (significant at .01) and Vocabulary (significant at .05). Digit Span, which is significant for Department A, does not hold up in Department B, and Picture Arrangement, which is

<sup>5</sup>It was found that the *MAT* correlates +.77 with the *W-B* Vocabulary subtest. Its next highest correlation is with the Information subtest. A paper reporting the results of correlating each of the *W-B* subtests with the *MAT* is in preparation.

TABLE 5  
CORRELATIONS BETWEEN THE *W-B* SUBTESTS AND CREATIVITY

Group	N	Inf.	Com.	D.Sp.	Arith.	Sim.	Voc.	P.A.	P.C.	B.D.	O.A.	D.Sym.
Consolidated	26	.38	.01	.48*	-.11	.15	.55**	-.12	.29	.21	.23	-.05
Dept. A	37	.28	-.01	.24	.29	.30	.20	.47**	.08	.29	.22	.09
Dept. B	63	.31*	.00	.32*	.13	.21	.31*	.29*	.14	.26*	.22	.04
Combined A & B												
Ph.D.s	9	.30	.61	.25	-.37	-.52	.70*	.42	.53	.05	-.04	-.04
Dept. A	15	-.02	-.18	.23	.65**	.50	-.38	.16	.12	.20	.28	.02
Dept. B	24	.06	.00	.26	.31	.25	-.11	.23	.16	.16	.22	.01
Combined A & B												
Non-Ph.D.s	17	.48	-.22	.61**	-.22	.34	.48	-.17	.21	.01	-.05	-.15
Dept. A	22	.48*	-.22	.07	.09	.06	.35	.49**	.33	.30	.27	.01
Dept. B	39	.47**	-.22	.29	.04	.19	.39*	.27	.24	.16	.13	-.06
Combined A & B												

\*Significant at the .05 level.

\*\*Significant at the .01 level.



significant for Department *B*, does not obtain for Department *A*. Thus, for non-Ph.D.s, it is the verbal<sup>6</sup> factor that consistently differentiates the more from the less creative subject. This finding is consistent with the results in Table 3.

#### F. DISCUSSION

Although the "consolidated" results reveal a significant relationship between intelligence and creativity, this relationship is confounded by the fact that education (Ph.D. or non-Ph.D.) was significantly related to both creativity and intelligence. After controlling for the effects of education, the results indicate that the intelligence tests used in this study do discriminate the more from the less creative non-Ph.D.s but do not discriminate among Ph.D.s.

What accounts for this difference in results? One might suspect that the failure to obtain a significant correlation for Ph.D.s resulted from too narrow a range of intelligence test scores. This hypothesis may be rejected on the basis of the data presented in Table 4 where the sigmas for both Ph.D.s and non-Ph.D.s are not different from each other.

To understand the different results obtained for the two populations discussed in this paper, it is necessary to recall that the subjects are now gainfully employed in research activities. While on the face of it "research activities" may imply that all men have equal opportunity for creative work, the reality of the situation is that such is not necessarily the case. Men are selected for their jobs and given opportunity for creative work on the basis of the expectations that the "significant others" have for them. It is suggested that for the population in this study different criteria were used for providing the men with this opportunity. For the Ph.D.s these expectations are based on a well-defined status symbol, the Ph.D. degree. Because a man has this degree, it is assumed that he has the intellectual ability and capacity to carry out independent research. Consequently, his intellectual ability is not scrutinized carefully. With the non-Ph.D.s, lack of the status symbol does not provide the "significant other" with a clear basis for developing expectations. He has to rely on other cues and, particularly for men with little experience, these are obtained chiefly from the interaction between interviewer and applicant in the pre-job situation. Thus it was hypothesized and borne out by interviews with the directors that the selector made a much more careful evaluation of the general intellectual ability of non-Ph.D.s than of Ph.D.s. Of the non-Ph.D.s, those men who made a

<sup>6</sup>See Footnote 5.

good impression were more likely to be placed in jobs having good opportunity, if such were available, as compared to those men not giving this impression. The latter were either not hired or were placed in jobs which provided less opportunity for creative work (more routine in character). It is apparent that verbal facility would be an important factor in differentiating these two groups.

To summarize the above: men with high status (Ph.D.s) are not evaluated for intelligence and, therefore, for them measures of intelligence are not related to creativity. For the low status subjects (non-Ph.D.s), high verbal intelligence raises their status and provides them with the opportunity for creative work. Therefore, for non-Ph.D.s intelligence *appears* to be related to creativity, but its primary function is to obtain the opportunity for creative work.

To check the hypothesis of the relationship between opportunity and creativity, an independent set of data was utilized. All the sections<sup>7</sup> in each department were ranked by the department heads for importance<sup>8</sup> of the sections' activities to the company. The ranking each man obtained for this variable was the ranking given to the section of which he was a member. It was assumed that an "important" section would be one for which the company would be willing to spend more funds and willingness to do so would remove some of the environmental obstacles that would limit creative work. The more and less creative Ph.D.s and non-Ph.D.s were then compared on the variable of importance. The results indicated that for Ph.D.s importance of the section with which a man was associated was not related to creativity. On the other hand, the more creative non-Ph.D.s were significantly more often associated with a section rated high on importance to the company than the less creative non-Ph.D.s (.05 level). This supports the hypothesis suggested previously that for non-Ph.D.s intelligence is a variable related to opportunity for creative work. Finally, the *IQ*'s for non-Ph.D.s were re-evaluated, controlling as much as possible for the effects of opportunity. All those subjects whose sections were rated high on opportunity were divided into a high and low creative group based on the original criterion ranking and compared for *IQ* scores. Neither the *MAT* nor the Full-Scale *W-B* differentiated these two groups. Consequently, when op-

<sup>7</sup>A section is an administrative unit consisting of men who work in the same technical area, although they do not necessarily work on the same problem. The number of men in a section varies from two to 10.

<sup>8</sup>It is to be pointed out that this criterion is not to be equated with opportunity. It is at best a first approximation and would be one of the factors that would be involved in a more thorough study of opportunity.

portunity is controlled for both Ph.D.s and non-Ph.D.s, intelligence is not related to creativity. The discussion further suggests the hypothesis that where opportunity for creative work is more or less equally available to a group of research workers, "*IQ*-ness" beyond a certain point (in terms of our population approximately the 95th percentile) becomes less significant for creative work. At this level other factors, such as emotional and motivational ones, far outweigh the importance of additional *IQ* points. Data relating these other factors to creativity are now in preparation and will be reported on in the future.

Finally, from a methodological point of view this paper calls attention to the importance of studying environmental factors in personality and assessment research. If attention is paid only to the individual in the equation  $B = fPE$  and the environment is overlooked, then significant relationships may be obscured.

### G. SUMMARY AND CONCLUSIONS

This paper concerned itself with a study of the relationship between measures of intelligence and creativity in 64 scientists actively engaged in industrial research. The creativity criterion was the rankings of superiors. The Wechsler-Bellevue and the Miller Analogies Test were the measures of intelligence. Since subjects with Ph.D. degrees were judged as significantly more creative than those without such degrees, the data were presented separately for both groups, as well as for both groups considered together.

The results of the study were:

1. Age, length of service in the organization, and total professional experience were *not* related to creativity.
2. The consolidated analysis revealed a significant relationship between intelligence and creativity. When education was held constant, however, this relationship did *not* obtain for Ph.D.s but *did* for the non-Ph.D.s. For the latter, the discriminating tests were the *MAT* and the Vocabulary and Information subtests of the Wechsler-Bellevue.

In the discussion it was suggested that men with high status (Ph.D.) are not evaluated for intelligence, and therefore for them measures of intelligence are not related to creativity. For the low status subjects (non-Ph.D.s), high verbal intelligence raises their status and provides them with opportunity for creative work. Furthermore, it was hypothesized that where equal opportunity is available, higher *IQ* scores beyond a certain point (approximately at the 95th percentile) have relatively little significance for creative work as compared to personality and social factors. The relationship between the

(18), Laszlo (12), Condrau (5), Becker (3), Savage (19), Busch and Johnson (4), Fervers (7), Frederking (9), Forrer and Goldner (8), Abramson *et al.* (2), Sandison (14), Sandison, Spencer, and Whitelaw (15), and Sloane and Doust (16) may be consulted. In addition, the papers of DeShon, Rinkel, and Solomon (6), Hoff and Arnold (10), and Rinkel *et al.* (13) are of importance. Sandison, Spencer, and Whitelaw publish a list of instructions to the nursing staff attending patients receiving treatment by LSD-25. They generally dealt with a different group of patients than those treated by the writer and the special patient reported here, in detail. Their instructions may well be used as a basis for therapy with an in-patient service. In the studies by the writer, the prolonged time of the interview or experimental period, often included discussions on the way home with a patient, at dinner, or wherever the patient happened to be. It is believed that the ambulatory technique not recommended by some writers is perfectly safe, if carried out with the dosage recommended and with the precautions mentioned. Oral administration was employed.

There is one important difference between the methods of Sandison, Spencer, and Whitelaw and the writer. They state: "No special preparation is required. The patient should be told at least the day before that treatment is to be given." I believe that the amount of anxiety which may be mobilized in ambulatory patients who expect to take LSD-25 is much reduced if a two-week period of notice and questioning is permitted to precede the LSD-25 interview.

### C. PREPARATION AND MANAGEMENT OF THE AMBULATORY PATIENT

In preparatory discussion with the patients, emphasis is placed on the therapeutic value of the drug. The unusual vegetative and psychic effects are related to the patient at least two weeks in advance of the interview with LSD-25. Anxiety may arise in spite of the most careful preparation. This anxiety should be met with confidence by the therapist who should inform the patient that the unusual psychic effects of the drug can only lead to beneficial results. The interviews may take place on the couch, sitting up, or with the patient walking about. Any special procedure usually practiced by the therapist in the interview situation should be replaced by a permissive expectation that the integrative functions of the patient's psyche will know best how to regulate the interview, provided that the patient remains within the *direct* control of the therapist. With doses of 40 gamma and below, the interview may last four hours but the effect of the drug may persist for another four to eight hours. The patient should be informed: (a) That



there are no after-effects. (b) That a relative or friend should call for him at a stated time (four hours after administration). (c) That he must remain at home or with friends until the next morning. The patient must be impressed with this *modus operandi*. While the effect may outwardly be negligible in certain ambulatory patients, it is the action of choice to keep the patient accompanied at home until the next day when all pharmacological effects due to the drug have invariably disappeared. (d) Driving a car, using machinery, or walking alone is prohibited. (e) A barbiturate, secobarbital ( $1\frac{1}{2}$  grams to 3 grams), is given to the patient to take at bedtime. (f) The four-hour interview is interrupted at about 50-minute intervals. The therapist is advised to have his secretary offer the patient coffee and simple foods throughout the rest periods. This may not only be desirable physiologically but also provides a milieu which diminishes the under-current of anxiety provoked by the drug itself.

For short interviews of one to two hours, advantage should be taken of the data provided in the first paper of this series. A study of the curves for doses of 50 gamma indicates that the peak reactions are observed between  $1\frac{1}{2}$  and 3 hours. This should be the period of choice. However, the patients have always been kept in the office under control a minimum of four hours. No method (except deep sedation) is known to completely stop the LSD-25 reaction once it has begun. Sedation may, however, complicate the situation more. One patient (F.W.) who received 50 gamma at 8:00 P.M. showed marked signs of the LSD-25 reaction at 1:00 A.M., even though 3 grains of secobarbital had been given at midnight. The patient was showing the signs of both drugs: elation (LSD-25) and an unusual lack of motor and sensory control (secobarbital).

#### D. CASE REPORT

A 40-year-old married woman, who had come for psychoanalysis because of many psychosomatic and adaptive difficulties had made excellent progress (1). For example, relationships with her two children were markedly improved. Her ability to work had been restored with elimination of nearly all of her psychosomatic complaints, such as eczema, asthma, hayfever, headaches, and similar complaints. However, there were still some important problems, especially connected with her sex life within her marital situation which required attention, even though vaginal orgasm was frequently achieved with her husband. These sexual problems were mobilized by a dream about two dogs preparing for the sex act. Analysis of the dream was unsuccessful because the patient's anxieties overcame her usual interest in

ascertaining the nature of the symbolic meaning of the dream. She said, "I am afraid of analyzing this dream." She had on previous occasions utilized LSD-25 to reorganize her defenses, so that in interviews lasting four hours, sufficient time permitted her both to reduce the defensive mechanisms and to reconstruct what was learned thereby to her psychic benefit. In the 50-minute interview ordinarily used, the psychic organization of this patient and perhaps of others is often left in a state more threatening to the patient than the patient can anticipate facing until the next interview! The patient soon learns to guard against this unprotected exposure. It is conceivable that with certain patients, particularly of the type presented here, that an analysis might take half the number of total hours if the interview times were doubled. In view of the satisfactory experience this patient had previously had in our interviews under LSD-25, it was decided to attempt to determine just what conscious and unconscious factors might be producing anxiety and the fear of analyzing the dog dream. In the following interview the nature of the anxieties connected with the dog dream were successfully analyzed by the patient, and a fear of homosexuality in herself which had been brewing but which had become intense was resolved. The verbatim recording which now follows outlines clearly a good part of the patient's psychosexual development which led to her confusing her childhood manner of relieving anxiety through masturbatory channels with later aspects of the psychosexual development, which both in the preanalytic and analytic frame of reference led to the patient's believing that she was a lesbian.

*Time: 2:00 p.m.*

*Dr.* Now you've taken a dose of LSD-25 by mouth. You took it at half past one, about 30 minutes ago, 4/10 of an ampule of LSD-25 and then went out for lunch. What did you have for lunch?

*Pt.* I had a tunafish sandwich and a glass of iced tea.

*Dr.* How would you like to start?

*Pt.* Well, I don't know if this should go down at all on the record, but I think I somehow or other, for some reason, I don't think it's because I anticipated taking LSD, but more of what I would talk about. I felt rather concerned all week. I was rather busy for a few days and those few days I wasn't able to have time to think about it. I didn't seem this concerned, but I wasn't too busy. I felt concern about it. And I also, yesterday, scratched my arm once, and of course was very aware of it, and even though it kept itching me I didn't scratch, and found it went away. I found that I scratched my back mostly in my sleep. I was pretty aware, also, of the fact during the week that although I knew I had dreams, I didn't remember the dreams. I really believe that I'm not a good subject that has, although I've discussed it here a good deal, I think anyway, perhaps not so much, but from

the beginning of the analysis it has always been difficult for me to discuss sex in any way, about myself or anything else. When I did discuss it, or if I anticipated the fact that I would discuss it at a session, I was always very tense before and afterwards. Although I realize that after all this time I, well, I am able to have sexual relations and that I am able to enjoy them more than I did, there are some times that I don't, or rather, not that I don't so much enjoy the relationship itself, but think about it seriously to a great extent. The one part about sex that still bothers me a great deal is masturbation. (Pause) What concerns me first of all is that very often, very often isn't the word, every time that John (husband) would go away I would find that I have the desire to masturbate. It used to be much more than what it is now, in that there would be many times when I would be home during the day, when before I, oh, I would take a bath or a shower, or something, well, anyway when I was going to take a shower, that I would very often practice masturbation. And I felt that even though I had sexual relations and enjoying them, that I still have this desire. It disturbs me very much. I was reading a book last night, a book called *The Sec-ond Sex*, and in it I, there was one chapter that was called lesbian, and I read that because even though we had discussed a good deal of this here, and you had said to me that you felt that I had enough feminine characteristics not to be considered a lesbian, I think I still have the desire, the masturbation left a question in my mind.

Dr. What question?

Pt. Well, whether I had some characteristics, I guess you could say, of a lesbian.

Dr. What do you mean by lesbian?

Pt. Well, now that you ask me the question, I realize that probably my idea, my conception of a lesbian is wrong. To me it was always a matter of, to the extent that I masturbate.

*(It is of interest to point out that the patient's psychosomatic complaints of eczema, asthma, hayfever, and other similar symptoms had been essentially eliminated, scratching now occurring in connection with the mobilization of anxiety over the fear of homosexuality.)*

Dr. Yes, but I didn't know that that was lesbianism.

Pt. Well, as I said, now that you ask me the question—

Dr. You wouldn't call boys who masturbate homosexual, would you?

Pt. No.

Dr. But that's what you're doing, when you talk about yourself. You see, I'm not saying that you shouldn't go ahead and discuss your anxieties about your masturbation, but I am emphasizing, "Why do you call that lesbianism?"

Pt. I said part of the reason was that in going back I think I always had the idea that sexual relationship with a man would be so satisfying that in time, or rather, after marriage, your desire to masturbate would die down. Yet it hasn't with me. I think another thing that

has to do with this, when John and I first started having relationships before we were married, the only time that I ever reached an orgasm was through manipulation. And even though, afterwards he would enter the vagina, I never really reached an orgasm. And then, in time, I used to think, well, I used to question, how do you really feel when you reach an orgasm? Could I think that I did, and didn't know that I had. And then I became concerned because I only seemed to reach an orgasm in one way, and so I felt that although it was with a man, and then in time, with my husband it still, well, it was still the same practice that lesbians use, which is part of the reason that, well, which is the real reason that I was concerned about it. Then, of course when our home life became so strained, and I came to the point where I was using sexual relations as a weapon, when I was so unhappy that I had no desire for them at all, that when I subjected myself to him it was such a strain that I became more concerned about it. As I started saying before, I started reading this book last night, *The Second Sex*, and well, I was rather sleepy when I was reading it. And somehow I just sort of remember this from it, that perhaps it does relate to me, in that it said something about a girl never getting past the clitoris stage, so to speak, where she does derive some sensation, some satisfaction, except by manipulating the clitoris, and of course I probably had no right to read it as I did. It also went on in some way as to the mental reasons as to why a girl falls into well, into lesbianism, and it talked about the relationship between the mother and the daughter in two ways that the mother can be overpowering, or that she can be a bad mother. That's very strange.

*[The patient's statement, "That's very strange," refers to the effect of the LSD-25 which was taken approximately 45 minutes before. The peak of the reaction usually occurs between the second and third hour, but depends upon the symptoms studied. See the first paper of this series (2).]*

Dr. You feel very strange?

Pt. Well—

Dr. How? Do you like it?

Pt. Yes.

Dr. It was all right for you to read that book. I think you could read anything you want to now. You're going to read a great deal, and you might as well discuss what you read with me.

Pt. When I read the book, I sort of well, I saw myself as the daughter with the dominating mother, and I'm not getting away from it. I'm afraid I'm going to have to reread whatever I read about that because I don't remember, or perhaps I have to read back to get the full idea of what was in the book. It seems also, that there are different situations which happened when I was younger, in that either I recall at times, and which I remember now. The first seems to be—I don't know if it's so or not—it seems that I remember the first time I ever had a sen-



sation of water on the clitoris, and how I discovered that it was a pleasant sensation. I can still picture the bathtub, I think, and the faucet, and how the water came out. And then I, of course, I recall one time when I turned around and found my father watching me; and I was so concerned as to whether he would go back and say something to my mother. And then, also I had such feeling about masturbation. I even became very superstitious about it, in that I was afraid to masturbate on a night before I was going to have an exam or something. The fear that it would, well, that God would punish me with different things, that something bad would happen to me that day!

*Dr.* Where do you think you got those ideas from? About punishment because you had sexual pleasure?

*(This is the earliest recollection of a guilt feeling in connection with self-stimulation.)*

*Pt.* I don't know where I got them from. I could only say, possibly, from my parents, probably my mother, because my father never discussed sex. And yet, I can't ever remember my mother discussing sex either. I recall, too, that when I was younger, my sister and I always shared a room together. I used to think it was awful if she walked around undressed. I just sort of remember that now. I remember my mother even mentioning it years after. How I used to think it was so terrible for my sister to walk around nude! This, of course, was in our own room. I found, I remember also, that when I said that I didn't like, that I thought it was awful her walking around nude, I also found, if I remember correctly, to see her undressed was repulsive to me, and yet I didn't have the same feelings about my mother. I would often be in her room when she got dressed. And then it was my sister who explained the menstrual cycle to me and although she didn't explain it clearly I think I somewhat knew about it, because I had seen her walk around undressed. I think I was pretty annoyed with her afterwards, too. I was going to say, probably most young girls want to grow up and be like their older sisters, and feel that when they get to the point of menstruation that they're pretty grown up, too. But I didn't have those sort of feelings. *(Long pause)* To get back to my relationship with John, I had, oh, I guess you could say, a good deal of involved thoughts and feelings about it. I was just trying to remember. I wondered if, I think it was more or less, trying to find an answer of whether it was I who wasn't able to have real sexual satisfaction, or was it because of John that I couldn't. *(Long pause)* I don't know. Either I'm back again to the question of why do I think about having sexual relations previous to having them. That is, why do I think—not always—every so often, why do I have to think during an evening whether when I go to bed he'll want to have sexual relations or not. That's the main question. Actually, I'll end it there. Whether he's going to bother me about having sexual relations.

*Dr.* But I understood that about your eightieth session you began

to have orgasms when your husband's penis was in the vagina. Is that correct?

*Pt.* Yes, but I would say, probably to begin with, it wasn't all the time.

*Dr.* But it would happen?

*Pt.* Yes.

*Dr.* What were your feelings then?

*Pt.* I was very happy; I felt more satisfied. I should say more satisfied with myself.

*Dr.* Well, what interrupted that?

*Pt.* Well, I think one thing that bothered me a great deal, the fact that when we were living in Boston, we got into a discussion with the son of my neighbor, and they sort of implied, and which I think, I gathered from what they said, that they could tell, or they could hear, every time John and I went to bed together, and every time we had sexual relations. I was very much aware of the fact that the people underneath us could hear us, and there again that's something that has always bothered me a great deal, in that nobody must know when we have sexual relations. It must be as private as it possibly can be.

*Dr.* Why do you think that it must be as private as it possibly can be?

*Pt.* Why do I think I feel like that? Or, why do I think it shouldn't—

*Dr.* Oh, no! Why do you think you feel that way?

*Pt.* The first thing I think about is the fact that I never knew, never once did I ever know that my parents ever had sexual relations. In fact, if I can remember, it was around the time when my father first took sick, I think, my mother mentioned something about my father not having any sexual relations because it was too much of a strain on him, something like that. Which was the first time, that ever, that they ever had sexual relations, was mentioned to me. And then the time when my sister started telling me that she had told my mother what had gone on between my sister and her husband, my reaction at that time was, what right have you to discuss that with mother? First of all, my feeling was, that she wasn't the one to discuss it with because her views were so distorted. By that time I was able to realize that. But even then I had the feeling that not only shouldn't anyone ever know when you have sexual relations, but you shouldn't discuss them with anybody. There was a time too, when I was also concerned about Ann, when I first noticed that she masturbated as a child, and I knew not to say anything to her, or draw her attention to it. I was pretty concerned about it, because I thought of it in terms of myself. Would this go on, as it went on with me? And look how I am today, and such feelings as that!

*Dr.* You said that when you had sexual pleasure through masturbation, you were afraid you'd be punished. You didn't know why. Now, as a little girl of eight or nine you hardly knew the meaning of sexual intercourse. You were wholly involved in sexual pleasure of self-stimu-

lation. You had no real knowledge of the mechanics of sexual pleasure, but yet you were afraid that the Lord would punish you by making you fail! And you feel that no one must know, not the slightest implication, even the neighbors that you were having sexual relations with your husband, even though this was none of their concern, none of their business. Does that bring up anything? It might throw some light about your concern, about your husband approaching you. Do you think that the same fear that you had as a child, that you'd be punished, has left you?

*Pt.* No, I would say that probably the feeling hasn't left me entirely, and especially I would say, the times when John approaches me, and it, it—especially at certain times, I think I still have those guilty feelings.

*Dr.* They're guilty fears, aren't they? They're not alone guilty feelings, but they're guilty fears.

*Pt.* Yes, they probably are the fears I had when I was younger, and yet it's hardly why. Those are the times when John and I, or rather I just can't explain to him, but when I feel, when there are tense situations. Let's say at the beginning of the year, when I was concerned about whether we would be able to get enough customers and such, and was worried. It interferes with my desire for sexual relations. With John, that isn't an answer. One thing has nothing to do with the other. I think he can't understand how my worries about work can interfere with my sexual desires.

*Dr.* But at the same time remember, you were always worried about failing at school. Are you carrying through your mother's ambitions for you at school? Are we still dealing with a fear of your mother, rather than a fear of your getting a sufficient amount of work to do?

*Pt.* Well, I don't know how much it has to do with my mother. Of course, it's—anyone wants to be successful. It's just a fear of not being successful.

*Dr.* Yes, but those fears were put into you by your mother if we go back over your early years and see that that entered into your relationship with your husband. Not for you to be successful as a wife and a mother was the problem, by your mother's standard, but for you to be *as* successful as your husband. Why should he take the courses? Why should he attempt to lord over me? Whether his success was going to be reflected in your future happiness wasn't the point. The point was that you were first, according to your mother, who should solve the problem of your own security, if you remember. And part of that solution is cutting a man to size.

*Pt.* Then I had a husband who turned around and said to me he didn't want a housewife, he wanted a business partner. He wanted someone working with him.

*Dr.* But at the same time he never rejected you in bed, did he?

Pt. No.

Dr. So, he wanted a woman in bed, but also someone who could function with him in a joint business. But as far as his own sexual instincts toward you went, they were certainly very active, even though you were helping in his work, or what turns out to be your joint work. He didn't only want you as a business partner, did he?

Pt. No.

Dr. So, I have the feeling also that you still have a fear of punishment that's coming in. That you don't succeed in what you want to do. I'd like to go back. I'm a bit confused. At one time I thought that your sexual feelings had improved considerably with John.

Pt. They have, and you see, the way I know that sometimes they're greater than others, or rather, that they should be, well, that they could be more so; there are times when sexual relations with John are far more satisfying. The orgasm I reach is far greater than at times when I'm not as sexually stimulated.

Dr. Women fluctuate a good deal. Some women are more sexually excited before the menstrual period, some during the menstrual period. Some require the absolute privacy which you just mentioned. Well, what would you say was the ideal situation in sexual relations with John? When you had all the sexual excitation which would be normal for you?

*(At this point, the patient is beginning to be able to speak freely of the details of the sex act. This is undoubtedly due to the action of LSD-25 in a therapeutic situation. More than 300 interviews without LSD-25 may be considered as "control" interviews. A four-hour session without LSD-25 is not available, but from the writer's experience, it is doubtful if this material could have been obtained in so simple and direct a fashion.)*

Pt. There isn't any one time, that I remember in particular. I mean, I would say that I would remember a few different times when we, when I was excited sexually to the greatest that I've ever been. It couldn't have been much more; but you see, I think what bothers me more about it is that I would say I could probably number them, and one shouldn't be able to number them.

Dr. What happened at those times?

Pt. Well, at those times, I would say, first of all, that the duration of the act itself was much longer, and there was no desire on my part to end it. While we were having intercourse it was pleasurable the whole time. Well, instead of having the desire to have it end, as sometimes I do, I had the desire to go on indefinitely. And then, at the end of the act, I would feel exhausted, and satisfied, tired and relaxed. But then there are times when I can't wait until it's over. I don't want to begin. And I can't wait until it's over. Rather, I can't wait until he reaches an orgasm. And then I think, I don't know, I don't know if it's because it takes too long to reach an orgasm, or what, but I



think there were times when I hardly could, before I have an orgasm, so that when he's finished, I don't feel completely satisfied. And then, you see, I'm afraid to tell him. First of all, I used the word *afraid* to tell him. I guess I am afraid to tell him that I didn't reach an orgasm. I used my childish masturbation for sexual satisfaction. And after this continued for a while, was when I started to question, what is an orgasm? Am I capable of reaching an orgasm, with the penis in the vagina. It got to a point that I couldn't reach an orgasm that way.

*Dr.* I understand that you did, later on in your analysis.

*Pt.* I'm talking about when I was first married. I think that fact that I still, well, that I still desired to have him manipulate the clitoris, I still had a desire, although there are times now when I don't. That the desire is still there still bothers me, because to me, it's that I haven't really grown up, that I'm not able to have sexual satisfaction or relations in an "animal" way, that I still feel a child's way of masturbating—

*Dr.* Is that lesbianism?

*Pt.* Well, having the desire made me feel that it was.

(Rest Period—ten minutes)

*Dr.* Let's see; the time is now 3:15 and you've had LSD an hour and forty-five minutes ago. How do you feel?

*Pt.* I feel drunk, like when I get drunk from whiskey. Give me one drink or two drinks and I feel happy and gay and sleepy.

*Dr.* Are you aware of everything that you're talking about?

*Pt.* I know everything I'm saying; I know what I'm doing; and I probably will remember everything.

*Dr.* Can you think clearly?

*Pt.* Yes, I can think clearly.

*Dr.* One of the things I think you ought to talk about now is about the great, well, say fear, you had in discussing your sex life with me recently even though we had gone over a great many of these things before. I wonder if you'd go into that. What the nature of those worries were! Why you didn't want to talk about it; or why you were reluctant to talk of it. What do they mean to you? Why did you feel that way?

*Pt.* I think, probably, first of all I think I was afraid to talk about them is the fact that, well, I've had these fears about myself, I think, for a long time. Something I had kept to myself for so long, too, and not talk about with anybody. I never had anyone whom I could talk, turn to, and if I had a question about sex, get an answer. And yet, I remember before I started the analysis, that was one of the things, when I was thinking about the fact that I need psychiatric help, that here was somebody I could turn to and ask these questions and get an objective answer. And yet when the time came I couldn't do it. And even while I was thinking about getting psychiatric help I used to think that maybe there is a family doctor I could go to and ask questions

and find out about myself. When I, I never, I guess I never had enough confidence in any of them to do that. And then when I started the analysis it was easier for me to talk about my mother, I guess, and I found her easiest to talk about first. And then I think I got to a point where I not only had the fears about myself, but a sexual point where I was afraid to find out what I feared was true.

Dr. Were you also possibly afraid that I would think less of you, or that I would punish you in some way or other?

*[Early in the patient's analysis, her method of adaptation to the therapist was similar to that which she employed towards her mother. This turned out to be an advantage (transference neurosis) in the analytic procedure.]*

Pt. Well, that you thought less of me, yes. Especially in the beginning.

Dr. Remember, as a little girl you were afraid of being punished for having sexual pleasure? And as a little girl, your fears were not connected with being a lesbian, whatever that may be to you. But it was a fear of punishment! And you think your concern about talking to me about all of these things were not only concerned with yourself, but that in some way I would be a punishing parent.

Pt. In a way you assumed a punishing rôle, yes, because, well, it was more or less like the rôle the teacher had when you're younger. Or in some way it looked to my mother, that you always had to be good, always. Well, by discussing these things with you I would, well, lower myself and so by lowering myself and your looking down on me would be a punishment to me, because I was always looking to be the best in the class with the teacher, the favorite. I remember when I was being confirmed, and I guess, there again, too, was a great deal of involvement because I remember the time, through that whole year of being trained for confirmation and everything, I strived a great deal to be the best in the class. I guess for two reasons at that time. First of all, because I wanted a certain part, a certain speech in the confirmation. My sister had that speech, and I wanted it too. Then also, I don't think this is too soon after the business I had at Sunday School where I lied about the teacher and I was so very much afraid of that, and it wasn't, I can't remember exactly, it was Dr. ——— who came up to the house that time, and the teacher, and it was more or less my being the best in the class that time, and redeeming myself too, to prove to him that I was very smart, I was very good. It was also at this time that I had a great deal of confidence in my social relationship with boys, and I had a great deal of trouble with my mother at that time. I remember that was around the first time that I started getting eczema on my forehead. I had a great deal of pressure from my mother at that time, about having dates, this one being popular, and that one being popular, and my not going out. It was all through this time too, that many times I went out of the house and said I had

a date and didn't have one. Even if it was with a girl! It's very funny that I should have had the desire before I started the analysis to be able to come and discuss myself freely, I thought, and then when the time came I just couldn't do it. As I say, I never had any opportunity of discussing sex with anybody, or feeling that there was anyone I could turn to and ask questions and get an answer.

*Dr.* How much difficulty are you finding now, talking about it?

*(The beneficial effect of LSD-25 is now especially portrayed. Although in a sense, the patient is a special case in a psychoanalytic situation, I have seen the same type of production of material during brief therapy with LSD-25)*

*Pt.* Today, I don't feel any difficulty in talking about it. In fact when we had the break, and I started thinking that this was a big question to begin with, a big problem to begin with, and unless I come out of, well, unless something comes out of this today, that to not have the answers, or to not know where to go from here, as to my sexual relations, and my sexual force, actually, it's my feelings about it that, well, that I've had four years now, and after four years I certainly must be at a point where I am able to do something about this too.

*Dr.* Well, you just pointed out that that was one of the most difficult things for you to talk about. You also pointed out to me that instead of talking about the things you went to analysis for, you talked about your mother, that that was easy. The things that really troubled you, that really were on your mind, you very carefully avoided, because they would disturb you. Isn't that right?

*Pt.* Yes, that's right.

*Dr.* So would you say you've had four years of avoiding problems rather than four years of facing things?

*Pt.* In part. There are times when I deliberately, or rather force myself into discussing the parts that I found very distressful at the time, and yet I was able to do it. I think probably what bothers me more right now is the fact that after four years I still found it, and up until today, I guess you could say, so straining to come in and talk about it. And then I had a time where I really eagerly discussed. I think there have been times when I probably have spoken more or less freely, but not easily. It's always been such a strain to say everything that's been on my mind.

*(The patient, lying on the couch), became too anxious to speak freely and was asked to sit up.*

*Dr.* I'd like you to sit up, it's easier for you to speak sitting up. See if sitting up you can speak even more freely. You can walk up and down, or anything. For example, you mentioned last week that you want to take up the dog dream again. Should you take that up now, or should you wait?

*(The "dog dream": two dogs having sexual intercourse. The patient frankly said she could not freely discuss this dream in the preceding 50-minute interview without LSD-25.)*

*Pt.* I think it's a good time for me to take it up now because I'm pretty relaxed now and there's a great deal of feeling about that dream. When I think about it now it had, the dream, first of all when I first started thinking about it afterwards, I started, it started the problem again in my mind of, does this concern my feelings again about being a lesbian? And, you know, I feel silly saying it now, because as I talk about it more and more I see how foolish, well, it's more or less like perpetuating the first thing that I ever learned about lesbians as a child and this just continued in my mind, this conception, and that was it. Although I knew more about them, or rather had learned more about them in time, still this old conception remained with me, so when I thought about it in relation to myself, it was still there as it was when I was a child. It must have been a very great fear on my part of being a lesbian, in time, as I grew older, because probably I would say, that I didn't understand when I masturbated at first. To me it was just a pleasurable sensation. I didn't know it was masturbation. Nobody ever discussed masturbation with me, nobody ever said anything. And so I think probably perhaps it had something to do with the fact, as I remember now, that when I first learned about lesbianism I sort of connected the two, my masturbating and lesbianism, and the two remained the same in my mind.

*(The following vividly portrays how the little girl laid a psychological foundation for identifying masturbation with homosexuality.)*

*Dr.* Now I have an idea. Do you remember what happened before you started to masturbate with water? You saw something. And your fantasies when you masturbated followed a certain pattern which might hook it up with lesbianism. Can you link the two together? What were your fantasies before you masturbated? The things that excited you sexually?

*Pt.* Yes, I saw my mother taking a douche.

*Dr.* That's right.

*Pt.* I thought she was taking an enema.

*Dr.* That's right. You actually saw the tip going into the vagina, didn't you?

*Pt.* Yes.

*Dr.* And then—

*Pt.* I didn't realize it was the vagina at the time, though.

*Dr.* But nevertheless that's what you saw.

*Pt.* Yes.

*Dr.* And then you told me that you had certain fantasies that would get you excited sexually. And then you'd masturbate. Do you remember what those fantasies were?

*Pt.* Yes.

*Dr.* Tell me about them again and see if you can make a correlation.

*Pt.* It was—I don't remember.

*Dr.* Try to remember if you can. *(Pause)* Would you like me to tell you?



*Pt.* Yes.

*Dr.* Well, as I recall, you visualized something going into your own rectum.

*Pt.* Yes.

*Dr.* Remember that?

*Pt.* Yes, the thought of an enema would excite me.

*Dr.* Yes, that's right. The thought of an enema would excite you. So you talk about that. If you can visualize as a little girl your not knowing the difference between a vagina and a rectum, your not having a clear knowledge of the physiology, but just a knowledge that that area was an erotic area, an area where a great deal of sexual pleasure was mobilized, and you had seen your mother—

*Pt.* Who slammed the door in my face. . . .

*Dr.* Did she do that?

*Pt.* It was then, after that point, that I had so much feeling about it.

*Dr.* Yes. So this feeling was also a sexual feeling. Had your mother given you many enemas?

*Pt.* Not that I remember, no.

*Dr.* Did you visualize your mother giving you the enema when you got sexually excited?

*Pt.* I guess, yes.

*Dr.* It was your mother then, giving you this sexual pleasure. Maybe that is the correlation that made you think that this form of sexual pleasure was lesbianism, because it was sexual excitement produced by a woman. As you put it, your mother was the powerful one in the family. Could that be the origin of your misconception? The childish misconceptions which you held on to? Your sexual excitement was produced by a woman giving you an enema, by a woman putting something into you.

*Pt.* Yes, that's right. Those were my first feelings at all about—

*Dr.* Yes. So, that's what you define as lesbianism?

*Pt.* But then I remember I got to a point where my feelings, probably after this, because, I don't know how soon afterwards, but I remember that at one time, or at any time after that, that the thought of my mother giving me an enema, I didn't want, I was afraid, in a way. No, not afraid. It was just that I didn't want my mother to do it, for some reason or other.

*Dr.* Well, you can see the reason now. Because that was the way in which you got sexually stimulated. Is that correct?

*Pt.* Yes.

*Dr.* Well, now I think we've solved the puzzle of lesbianism. What do you think?

*(At the height of the LSD-25 reaction, note how the patient is able to integrate many facts, and simultaneously bring up data she previously was unable to discuss without overwhelming anxiety.)*

*Pt.* It's so funny, something like this should be, I mean, these in-

cidents. I can remember that I talked about them a long time ago, and each one of them was of individual interest in my mind, and yet there are only a few things that happened around that time and in that apartment that I remember. And there again, too, was the time when I thought my mother was gone and I thought something had happened to my mother, and I prayed to God, and I was crying so, all afternoon, because I was afraid something had happened to her. Undoubtedly I had a great deal of feeling at that time, and yet it took me all this time to talk about it, about them altogether, because, after that one afternoon when I said to you that I was always afraid of being punished by God when I masturbated, because after that afternoon, I remember I'd always felt that by getting down on my hands and knees and praying to God that he would help, because of that afternoon. He brought my mother back. And so I always have this feeling that he could punish me at these times, and also that he could be good to me.

*Dr.* Do you prefer sitting up or lying down?

*Pt.* I'm comfortable. I think I'd rather sit up. To get back to the dog dream—though, first of all, I thought of this before and I just recalled thinking it again, that I think one thing that I never had any satisfaction, or any feeling good about, when I used to dress in front of John, instead of feeling confident of the fact that I could, that I stimulated him in any—sexually, instead I used to feel kind of annoyed or embarrassed, and then, for so long, when I used to get undressed when I first came here, even, as soon as I started taking off my clothes I'd scratch so. Those were the times when I felt itchy. The dog dream, to get back again to that, it's somewhat of the same thing, somehow with that dream that even though in the dream I expected the dog to fight with the female, instead of really mounting her in the dream, he was licking her. And though I kept waiting for him to mount he never did. He just licked her. I said before that the questions that I had about John also just as an individual, I guess part of it, I guess one of the questions that I had—*(Pause)* That's another thing, too, because I guess, actually, when I think about how I derive my knowledge of sex, I couldn't tell you that I ever read a book, or that I did this or that. It's just, it seems where most of it came from, I don't know. Most of it even came out of my own mind. There wasn't too long ago that I, oh, I guess a couple of years before I was married, no, I don't even think that long, maybe a year before I was married, that I came across the word *cunnilingus* in a book, I think that's how you pronounce it, that I looked it up in the dictionary. I don't even, I don't know where I got the idea from but to me it was a, this is the wrong thing to do, too. I don't know why I think of this. Just about a week or two ago I thought of this too. I used to go through my mother's drawers when she was out, oh, I think up until the time when I left home, and in going through her drawers, I came across the books which I mentioned here before. They were, I guess, supposed to be

French books. One was a story which had lesbians in it too, I think. And the other was a book on the different positions to be taken during sexual intercourse. These two books were hidden in the bottom of the drawer, and I never mentioned it to my mother until, not even, oh, I think it could even have been after my father died, pretty recently though. And the other day I just realized, or I thought of it, was that when we cleared out her drawers those books were gone. My mother denied them at the time, said she never had them. It only struck me very funny at the time that she should deny it. I mean, at this late date.

Dr. Well, maybe she was embarrassed. Maybe she also had fears about sex.

Pt. Oh, I think my mother had a great deal of fear about sex. A great deal. Because I know that she had such distorted views. But from what she said to my sister, I know that her ideas were very strange. I probably derive some of them from her, I don't know.

Dr. By the way, it's about two and a half hours from the time you've taken the medicine. Do you still feel it, or is it wearing off? Do you feel as high?

Pt. Not quite as high. I'm getting more, (Laughter) I feel a slight headiness. Is it two and a half hours?

Dr. Yes. Why? Did the time seem much shorter?

Pt. Well, when I looked at the clock before, I just remembered now I didn't realize it was three o'clock, or that the time had passed so fast. I couldn't remember why time meant so much to me. Now I remember that the last time we talked about the fact that time goes much faster than what you realize.

(Time often appears to pass much faster while under the influence of LSD-25. This facilitates a four-hour interview for both patient and therapist.)

Dr. This time you implied that John was the dog in the dream. When we previously analyzed the dream we said that you were the dog in the dream. So it's just as if you had reservations about John's masculinity, which has worried you a good deal.

Pt. Yes.

Dr. Are you worried that you're not very much of a female because you married John? Is that what's really bothering you?

Pt. You see, before I said something about are we perhaps suited for each other because he is not as masculine and I am not as feminine.

Dr. Yes, but do you think masculinity and femininity can be measured like a yard of cloth.

Pt. I guess I'd like to.

Dr. You see, you haven't ever shown any important desire to derive sexual pleasure from a woman, although in your childhood fantasies it was your mother with a symbolic penis who stimulated you.

Pt. Yet I felt very repulsive towards my mother.

*Dr.* Yes. Well, that could be a compensation for your sexual excitation produced by the all-powerful figure in your home, the mother. And your mother with an enema tip in her hand, symbolically, could have frightened you. Your mother was a frightening person. And that very sign of strength would appeal to the feminine core within you.

*Pt.* Yes, I just remember now, too, I remember playing with my dolls. I would play that I was giving them an enema.

*Dr.* You were placating your mother, symbolically; you were a good little girl. She expressed her masculinity not only towards your father, whom she regularly castrated, but also towards you, whom she treated as a female, in your own symbolic thinking as a little girl. Your rôle has been essentially feminine, although you've been trying to placate your mother, living up to her idea of what a feminine woman should be like; in other words, someone like herself. She really tried to mold you in her image.

*Pt.* Then, in other words, as I said, I always kept my childish conception of a lesbian, even though my knowledge about them increased with time, so that actually, what I have done is that somehow or other I have taken physical reactions and disregarded my mental ones, in my thinking, and disregarded that completely and taken only my physical reactions as being indications as to whether I am a lesbian or not.

*Dr.* What do you mean by physical reactions?

*Pt.* Well, as I said before, the only way John ever excited me when we first got married, was by manipulation, and to me that was just a physical reaction.

*Dr.* But it was not lesbianism, which is—

*Pt.* No, so I had taken purely my physical reactions as being a guide as to whether I am or not, and disregarded whether I ever had any inclinations or dreams or feelings towards a woman. Which I never had.

*Dr.* Now, in this dog dream, you're really the female, aren't you?

*Pt.* Yes. And I'm waiting to be mounted, that was my feeling.

*Dr.* Now, you've analyzed the dream. And yet you tell me that your husband is willing to do that whenever you wish.

*Pt.* Yes.

*Dr.* So?

*Pt.* Which has always put a question in my mind about him.

*Dr.* What question?

*Pt.* Well, in that he is always ready to, well, to satisfy me sexually in any way that I wish.

*Dr.* What's wrong with that?

*Pt.* Well, I thought it was wrong. As I think about it more and more, I feel it was wrong, or I thought it was wrong because I felt that what I wanted was wrong, that he wanted to do it was wrong, too.

*Dr.* Wrong? Punishment again involved? In other words, you're saying it was wrong to have pleasure, sexual pleasure. It might not have been the most mature relationship, but I don't see anything wrong.



*Pt.* You see, when you said just the two words, sexual pleasure, I thought also of, yes, I think I was afraid, and sometimes I am afraid to have sexual pleasure, because I think of the time when I went with my friend to see the sailors and my mother found out and she spit in my face and called me a dirty tramp. To me, to have a desire to see a sailor, it was sexy, it was dirty, it was wrong, and my mother spit in my face for it. It was a horrible thing, and so I, also the idea that I was oversexed. I had all sorts of fears.

*Dr.* Of fears, which your mother injected, which you don't need anymore.

*Pt.* But why, I mean, as I think now, the times when I have said that I had the greatest sexual pleasure were the times when I was probably least inhibited during the sex act. Then what causes me to, why do these fears at times come out more than others? Because of the influence of my emotional feelings at that time? Of different influences upon me?

*Dr.* Your analysis is concerned with metamorphosing your fears of your mother and her dominance over you, and your understanding of those fears, getting rid of them, together with the guilt.

*Pt.* And I've had more trouble lately, since I'm back again in her house. That's one thing.

*Dr.* Yes, that's one thing. Ever since you're living in her home your fears have been again mobilized. You want the reassurance of a strong character like your mother. In terms of your mother, your husband isn't a dominant male, because he does what you wish. He gives you sexual pleasure the way you want it. Yet you describe him as being pretty aggressive. And yet you can't accept him because of the threatening nature of his being superior to you. You're afraid of what would happen to you, that you will become a passive female. You're afraid then that you'll lose him, which your mother established when you were a little girl, "This is what I want my little girl to be, she's got to be like me." In your first session, or second session, you were afraid you had cut your husband to size, too.

*Pt.* Yes, there's a great deal of conflict in my feelings towards John, sexually, because in a way I want him to be the aggressor, and yet when he is the aggressor I am annoyed, also. (*Laughter*) You see, it's true, whatever I have said here as to my sexual relations with him, or in whatever way I've spoken about sex, I've always made it as honest as I possibly could at the time, or if I told you of an incident it was true, because I realize to lie would be foolish, I mean, I'm not gaining anything by it. Either don't talk about it at all, or tell the truth. And yet, I realize now, that I've been quite concerned; I think my concern has turned more to sex again, in my relationship with John, because it wasn't what it was before I moved into my mother's apartment. I realized it, and I guess you might say it gathered momentum until it reached a point, until I couldn't stand it any more.

Dr. That's why I thought it was a great mistake for you to live in your mother's apartment; to be with her things, with her clothes.

Pt. Well, you know, the tangible things are very easy to get over. And to tell you the truth, I just realized before when I was sitting in the chair, too, the dress I made over for my mother, although she never wore it, it was my mother's, that I made over. I can put on most of the clothes now and not think too much about it. Oh, I'll take out a pair of gloves once in a while and think that it was my mother's but, oh, the tangible things as I started saying, you can get over them much easier, you're aware of them, they're in front of you. You see, it's the intangible things that you don't see, and you just don't realize.

Dr. That's what psychoanalytic theory deals with. Feelings are intangible things. It's not the dress, but the feeling that you have about the dress.

[Compare with Kinsey's discussion (11, p. 170).]

Pt. I wonder now, what I think I was, what I thought I was going to conquer by staying there, but I thought that I had something to conquer by living in that apartment and living through—

Dr. You had *someone* to conquer, instead of *something*.

Pt. It wasn't my mother I wanted to conquer. It was either John or myself. And I think it was he. Because by conquering him I was sure of the results.

Dr. Well, that was your mother's method of operation, wasn't it?

Pt. Yes.

Dr. See, all of your discussion about John in the last few months have been tinted with the notion that *he* wanted to be successful. That you couldn't take. In other words, his successes were your failures.

Pt. And that he was successful without me. (Pause) Unfortunately, though, they're not pure feelings. I could say that there again, there was a conflict with him. Of course, I could say that there were times when I want him to be successful, because I realize that his success could be my success, too. Yet it was always the idea that with his success I must have a finger in it also, that it isn't purely his alone.

Dr. But he wants you to! But he gets the glory. Isn't that what bothers you?

Pt. Yes, that's what bothers me.

Dr. When you went to that Board of Trustees meeting, do you remember?

Pt. Yes.

Dr. He was moving into the upper echelon of policy-making and you weren't.

Pt. But, you see, I have more or less taken the step though, in that, more or less, I think unconsciously, I think, though I am sometimes aware of it, I am sometimes, well, he's taking his step that way, and I'm taking mine this way, for me, and he joins this organization, and I join that organization and I don't want to be with him, so to speak, and

a member of an organization with him. I, it must be all for him alone or all for me alone, because I can't share with him.

*Dr.* Why not?

*Pt.* I could answer you very simply and say that I, because I've always had the feeling that whatever my father was and made of himself my mother helped him. My mother was the one who pushed him into taking steps to further himself, along in the business world, let's say. I don't, you see, I haven't been able to gain that feeling yet with John. It still is this idea that I'm working for him, or that I have done so much and have not gotten the glory I want, that he gets the name, that he gets the recognition.

*Dr.* But that's precisely what your mother always told you that you must do. She always told you that you must look out for yourself. And you're still doing that.

*Pt.* Yes, I am.

*Dr.* You said about 20 minutes ago that you realize, realistically, that his success was also security for you.

*Pt.* Oh, yes, I can realize that. But emotionally—

*Dr.* Emotionally though—

*Pt.* I've got to have—

*Dr.* It's threatening to you.

*Pt.* It's threatening?

*Dr.* Yes. You say that—

*Pt.* Because he has what I want?

*Dr.* Yes. (Pause) It's now three hours and 10 minutes after you've had the medicine. How do you feel now?

*Pt.* I still feel lightheaded. Not as much as before.

*Dr.* Well, do you still feel you can talk about things?

*Pt.* Yes, just as free.

*Dr.* Now, you've been sitting up.

*Pt.* I can repeat some of the things I said that we missed.

*Dr.* All right, go right ahead.

*Pt.* While you were out I was thinking when I had the question that I had when I gave the conclusion to what I thought I accomplished in analysis, and when I asked the question, "Where do I go from here?," was that I knew that the one big problem that I came in with and that I had not really gotten to the point of discussing freely or easily yet, was sex. And that I was questioning what am I going to do about it, really.

*Dr.* Which is also borne out that in the sessions following much of the material was the discussion of your relationship with your husband. Do you recall that? So that's what I felt was the problem. You'll remember that those sessions are earmarked by that type of discussion. In what way does the medicine which you took make you speak more freely?

*Pt.* Well, I first of all, I feel no tension when I'm talking at all,

and I just go on thinking and whatever I think I feel free to say, which I never feel, usually, especially when I'm discussing sex; I don't feel free.

*Dr.* Your mother didn't feel free to discuss sex. She hid the books which you mentioned.

*Pt.* She hid everything. I mean, sex just wasn't a word that was said in the house, and when it was said it was a dirty word, it just wasn't used. It was either a sexy person, or a sexy book, or a sexy joke. It was used derogatorily, not in any way of speaking objectively or any other way about it.

*Dr.* Then, in a sense, I think another point is brought up. You were brought up to feel that if you're married to a man who is always ready to mount you, so to speak, it was a derogatory situation.

*Pt.* Yes. And you were just as derogatory by letting him do it.

*Dr.* Yes. Yes, and that might account in a very definite way for some of the feelings you have.

*Pt.* Yes, you see, there again she really put conflict in me too, I guess: well, she more or less told me that the wife submits to the husband sexually, but you control, in a way, the number of times, or in some way or other, the wife is the controlling, and yet submissive at the same time. You submit to sexual relations, but you're still the controlling factor of it. That is, you are the one who says what he does and what he doesn't do, more or less, which she even told my sister.

*Dr.* So the only way you have complete control of your sexual life is to masturbate.

*(Judging from Kinsey's report, it is hardly likely that the psychodynamic significance of this point is understood by Kinsey and his co-workers.)*

*Pt.* Yes. And yet I don't feel control over it, because, well, that's probably why I have so much feeling over it, because well, that—

*Dr.* When I say control, I mean control of your sexual life with your husband, not control of your sexual life.

*Pt.* Well, that's what I mean. I don't, it, I do think about the fact that if I masturbate it might kill my desire to have sexual relations with him. It also enters my mind too, that I'm satisfying myself alone, and am I going to be able to enjoy sexual relations with him. And that I don't want either. So, I have so many conflicts about it, about the problem that I guess I—

*Dr.* Well, now you see all these conflicts, and you see their origin. But it isn't enough to see them. You have to make use of the insight which you have.

*Pt.* I think I can, because I think of the little, or anyway, of what I have discussed here before I have been able to use. Probably if I had been able to speak more freely before about this, perhaps I would have gotten over this big hump sooner. Actually, by this discussion today too, I think first of all, just the fact that I'm able to discuss it



freely is one big thing to me. And then there have been many questions that have been answered today. Because I guess, I, by letting the origin of why I have such feelings or thoughts, and by understanding, I think the need for some of it will disappear too, in time. (*Pause*) I was just thinking of an incident when my mother was in the hospital the last time and she was pretty sick at that time. My sister was there, and somehow or other, as my sister said, I don't know how, sick as my mother was, she just sensed that my sister was having her menstrual period and said something to her about it and you know, when I think about it now, it just sort of, the same way that sex always was all my life, with her, and yet I know that my feelings and my thoughts or whatever I was thinking about sex, were gotten from my mother, the unspoken word, nothing has to be said, she just somehow got it across to you. (*Pause*) You know it's, before I came here today I knew too, that this was what I had to discuss here, part of what had to come out and, oh, I guess I was wondering away whether for four years I had sort of side-stepped the real question, the real basis of what I needed the analysis for. I think it's all part of the picture of my relationship with my mother. I think I did side-step it at times, but I had always side-stepped it all my life, I always had to. And I think by her dying, more or less, well, in a way, I was going to say at first it stopped me from talking and yet it didn't. As a fact, it really brought it to a point, a climax. I really had to do something, about the one question that was still bothering me so greatly. I have one question I must ask you. How can you remember after such a long time, such a simple point of the one thing I told you about, about walking into the bathroom on my mother when she was taking a douche?

*Dr.* How do I remember it?

*Pt.* Yes. I mean, I've said so much in four years. How can one point be remembered? At different times I have thought this also.

*Dr.* It was a very important point. It was important enough for me to remember, to keep in mind, in connection with your problem.

*Pt.* Oh! Because I know I talked about it a long time ago.

*Dr.* Yes. But you said following that incident you had begun to masturbate, and then about two years later you were able to discuss the fantasies you had in masturbating; and now today those fantasies became crystalized and their meaning became clear. But it was very important that we know, that we both kept in mind what had happened just before, in connection with those fantasies.

*Pt.* I didn't recall it in sequence though. I didn't remember the fact that I started masturbating after that. I don't think it meant as much to me at the time—

*Dr.* That's to be expected, that you wouldn't make those correlations.

*Pt.* I remember there was something else—

*Dr.* You brought up the fact in connection with the dog dream that when you took the dog to the veterinarian that you got some hayfever.

*Pt.* Oh, yes, that's what I wanted to say. Yes. I sat there for about, this was last Monday, I was here last Monday, I took the dog to the veterinarian in the afternoon and I was in there, oh, less than five minutes, and I started sneezing, oh, terribly!

*Dr.* Were there many other dogs around?

*Pt.* Oh, yes. There were dogs and cats there, too. I would say that (*laughter*) I was emotionally upset at the time, and I realize that I was. Or rather, I'd say that at first I didn't, that at first I blamed it purely on the fact that being at the vet's, so much dander and everything around, and I was pretty bad off. But then it continued after I left there, and the next day too, I think I was still sneezing a little bit, and by that time I realized also, I was rather concerned at the fact that well, that I was coming here and that I was going to discuss all that I had kept in me for so long, and I realize that the two together make a very good combination, so that I had very good physiological symptoms of being extremely allergic to animals. (*Laughter*)

*Dr.* You were giving the dog away, weren't you?

*Pt.* No, I just took him for an injection.

*Dr.* Oh, I see.

*Pt.* (*Laughter*) But I had a very difficult time explaining to the vet how I could live with a dog and be so allergic at the same time.

*Dr.* It could have been the cat dander. Or it could have been the insecticide. I quite agree with you that the emotional factors are important, but I also believe there was a good deal of dander around.

*Pt.* Oh, yes, but I mean, I would say I could have had some reaction to it, but I think it was because I was emotionally upset at the time too, that it had so much effect. My eyes were all swollen and tearing. My nose became all clogged up.

*Dr.* The only way we could find out is for you to go back and see if you get the same reaction.

*Pt.* That I could do.

*Dr.* If you're interested go ahead and do it.

*Pt.* I'd be too much aware though, I think, I mean it would be so purposeful.

*Dr.* Well—

*Pt.* Oh, I've been there—

*Dr.* Oh, you've been there before?

*Pt.* Oh, yes, I've been there before.

*Dr.* During the pollen season?

*Pt.* No, this was during the winter that I was there. To his place, that is.

*Dr.* You've never been to his place—

*Pt.* I've never been to his place during the pollen season, but I've been to the vet's during the pollen season.

*Dr.* Oh, you have been?

*Pt.* Yes.

*Dr.* Last week there wasn't too much pollen in the air; it was between tree and grass pollination.

*Pt.* So, I mean, I have been to the vet's before, during pollen season. I know I take the dog every spring for a rabies injection. And I never had this extreme reaction, that I could remember. There was one time when I came back and told you about it. Of course, I still get hives when I touch a dog.

*Dr.* I'm inclined to think that you—

*Pt.* I can't pet him—

*Dr.* Sensitivity to dander—

*Pt.* I mean, right where I touch him I just come right out on my arm with—

*Dr.* I agree that psychological factors are important; but still, it's surprising that you can live with a dog. You get hives when you come in contact with his saliva?

*Pt.* Even his skin, if I just pet him. His hair, I get the hives from that.

*Dr.* He must have saliva on his hair.

*Pt.* He could because he licks himself a good deal. I think it's more his saliva than his body.

*Dr.* Yes, I think so.

*Pt. (Laughing)* And yet I hate to talk about that because my mother always used to say that she was allergic to dogs.

*Dr.* Oh, your mother was allergic to dogs? Well, that changes it a little bit.

*Pt. (Laughing)* And she always started scratching after a dog licked her. I think, I mean, I think some of it was a little psychological, also, because she always scratched, but she couldn't stand the saliva of a dog.

*Dr.* Now, do you think we've gone over enough ground?

*Pt.* I think so.

*(This concluded the psychotherapeutic interview. The patient was accompanied home by a friend, and felt well and slept well the next day. Her ability to integrate effectively new analytic material persisted for at least 10 days. As mentioned in the foregoing, she lost her fear of homosexuality and the anxiety which had been mobilized by the dog dream.)*

## E. DISCUSSION

According to Kinsey and his co-workers, masturbation is a type of sexual activity by means of which the female most frequently reaches orgasm. Moreover, their data provide evidence that it is the second in order of frequency of sexual activity before and after marriage for the female. In the foregoing case report, the psychodynamic significance of the masturbatory act is portrayed in connection with the life of the patient as a whole. We

have seen that masturbation may produce an extremely complicated pattern in the life of a married woman who had also achieved satisfactory vaginal orgasms with her husband. But life with her husband was especially complicated by her masturbation. Masturbation, therefore, was not a source of sexual pleasure, but a quasi-autonomic technique of relieving the type of anxiety that the patient had experienced in childhood in relationship to her parents. Psychodynamically speaking, scratching and masturbation are here connected in the following way with scratching and eczema which began in infancy. The threatening relationship of the mother (parent) to the child retained, for all practical purposes, her infantile eczema into adult life. This eczema, together with the threatening relationship, began at an early age when the prototypes of later guilt feelings were developing in the child. The reaction of the child and adult was characterized by a simple and violent infantile reaction to the threatening parent by a retaliatory attack on the only object the infant can attack—*itself*. It is thus that the allergic adult with persistent infantile type of eczema meets certain threatening daily life situations; their threatening nature is matched by a persistent and infantile form of retaliation—scratching in areas already prepared by an allergic constitution.

The little girl grew up, therefore, with scratching, itching, and a skin condition as one of the devices which she used to maintain a less threatening relationship with her parents, especially with her masculine mother who "cut a man to size." As the little girl passed through the Oedipal situation there was no feminine woman with whom to identify, but only a threatening phallic mother. Her adaptation to the male at this period was characterized by contacts with the passive father who had been effectively "cut to size." Although data are not available, it is conceivable that accidental manipulation during bathing by the mother may have erotically stimulated the little girl so that she found an important source of pleasure in her relationship with her mother. Her special technique may have been produced by the accidental observation of the mother taking a douche. In childhood fantasy, sexual stimulation was achieved by thinking of the mother inserting something in the patient's rectum, such as giving an enema. This fantasy is symbolically resolved in the analytic material by achieving orgasm through masturbation in a special way—by water. The water streams out through the tap, but also symbolically speaking, through the enema tip held by the mother and thus participated in, in toto, by the phallicized mother or mother-symbols. It was only through psychoanalysis of this pattern as given in the verbatim record that the patient was able to eliminate the concomitant fear of this symbolic homosexual relationship in her daily life.



One wonders if this psychodynamic approach is not more important than that emphasized by Yerkes and Corner in their preface to Kinsey's book (11):

Comparison of Freud and Kinsey is not implied, for the two men differed greatly in temperament, professional training and experience, and in their objectives; but what should be noted is the fact that Freud, on the basis of his clinical experience, proposed theories which laid the foundation for a task he was not fitted by nature or training to carry on. This is the great task of fact-finding through careful, patient, long-continued, objective research which Alfred Kinsey, the laboratory- and field-trained biologist, is now engaged in doing. From the Kinsey project, sufficiently extended, should come basic knowledge of sexual phenomena against which theory may be checked, modified and supplemented.

That members of both sexes masturbate is well-known. However, the psychodynamics of any particular case is usually extremely difficult to elicit, and I believe that contrary to Yerkes and Corner, it is Freud's work which laid the basic foundation for studies of the type to be reported here, and that enumeration of data, however statistically valid, in the fashion employed by Kinsey and his group, with no data on unconscious motivation, with no understanding of the individual, but only of one minor fragment of that individual's behavior, will hardly lead to an understanding of the rôle of sexual behavior in our culture.

It has generally been recognized by the investigators mentioned in the historical part of this paper that LSD-25 produces a disturbance of the personality structure, this disturbance being analogous to a schizophrenic-like state in which ego-depression occurs. I believe that this generalization, although valid, must be modified. The effect of LSD-25 depends upon the dosage and upon the personality of the individual. Although ego-depression does occur in the normal individual taking up to 50 micrograms by mouth, I believe that one other process goes on simultaneously, and that process is ego-enhancement, or reinforcement. The ability of the individual under small doses of LSD-25 to face preconscious or unconscious material and to integrate this material into the dynamic forces of the ego-structure is not part of a loss of ego, but rather that of a reinforced ego which functions more effectively under the drug, *in the presence of the therapist*. A study of the psychoanalytic material of this particular patient after the LSD-25 interview reveals consistently that the experience under LSD-25 gives the patient much more confidence in reconstructing and reevaluating data for an extended period of weeks than had previously been possible without LSD-25. This integration usually required very extended periods of

therapy. I prefer to think of the process of integration under LSD-25 as distinguished from narcosynthesis, as hebesynthesis, an elated state in which both ego-depression and ego-enhancement may occur simultaneously with the ego-enhancement leading to an increase in the integrative functions of the patient's ego.

#### F. SUMMARY

1. Lysergic acid diethylamide (LSD-25) provides a relatively new adjuvant in psychotherapeutic procedures. The use of the drug is characterized by pharmacologic safety, effectiveness in small doses, maintenance of the patient in a conscious and coöperative state, and by safety on repetition without evidence of addiction in both ambulatory and hospitalized patients.

2. Directions are given for the preparation of ambulatory patients with special particulars for 4-hour psychotherapeutic interviews.

3. A typical 4-hour interview is presented verbatim. A woman of 40 became anxious during psychotherapy because she feared that she was homosexual. During the 4-hour interview under LSD-25 the psychodynamics of this fear of homosexuality was reconstructed in connection with her technique of and fantasy during masturbation with running tap water.

4. Under LSD-25 the integrative processes of the ego became more manifest and she was able to lose her fear of being homosexual. The data indicate that with a low dosage and the technique employed, LSD-25 not only produces the well-recognized schizophrenic-like state with ego-depression but also a simultaneous process occurs that may be termed ego-enhancement. This ego-enhancement leads to hebesynthesis, an elated state in which the processes of ego-reconstruction result in reinforcement of the integrative functions of the ego.

5. The statistical point of view of masturbation (Kinsey *et al.*) is compared with the value of studying the phenomena within the total configuration of the personality structure. It is held that the solution of the psychodynamic problem of the individual must be part of any realistic evaluation of the meaning of masturbation by the female in our culture.

#### REFERENCES

1. ABRAMSON, H. A. Technique for screening verbatim psychotherapeutic recordings and its application to allergy. *Ann. Allergy*, 1951, 9, 19-30.
2. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., LEVINE, A., WAGNER, M., & KORNITSKY, C. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psychol.*, 1955, 39, 3-60.
3. BECKER, A. M. Zur Psychopathologie der Lysergsäurediethylamidwirkung. *Wien. A. f. Nervenhk.*, 1949, 2 Jg, 402-440.

4. BUSCH, A. K., & JOHNSON, W. C. LSD as an aid in psychotherapy. *Dis. Nerv. System*, 1950, **11**, 2-4.
5. CONDRAU, G. Klinische Erfahrungen an Geisteskranken mit Lysergsäure-Diäthylamid. *Acta Psychiat. et Neurol.*, 1949, **24**, 9-32.
6. DESHON, H. J., RINKEL, M., & SOLOMON, H. C. Mental changes experimentally produced by LSD (d-Lysergic acid diethylamide tartrate). *Psychoanal. Quart.*, 1952, **26**, 33-53.
7. FEVERS, C. Die Narkoanalyse als initiale Methode in der Psychotherapie. München: Lehmann, 1951. Pp. 120.
8. FORRER, G. R., & GOLDNER, R. D. Experimental physiological studies with Lysergic Acid Diethylamide (LSD-25). *Arch. Neurol. & Psychiat.*, 1951, **65**, 581-588.
9. FREDERKING, W. Ueber die Verwendung von Rauschdrogen (Meskalin und Lysergsäurediäthylamid) in der Psychotherapie. *Psyche*, 1953, **7**, 342-364.
10. HOFF, A., & ARNOLD, O. H. Die Therapie der Schizophrenie. *Wien. klin. Wchnschr.*, 1954, **66**, 345-351.
11. KINSEY, A. C., POMEROY, W. B., MARTIN, C. E., & GEBHARD, P. H. Sexual Behavior in the Human Female. Philadelphia: Saunders, 1953. Pp. 842.
12. LASZLO, M. Meskalin- und Lysergsäurediäthylamidrausch. *Grenzgeb. der Neur.*, 1952, **12**, 146-147.
13. RINKEL, M., JACKSON, H., DESHON, H. J., HYDE, R. W., & SOLOMON, H. C. Experimental schizophrenia-like symptoms. *Am. J. Psychiat.*, 1952, **108**, 572-578.
14. SANDISON, R. A. Psychological aspects of the LSD treatment of the neuroses. *J. Ment. Sci.*, 1954, **100**, 508-515.
15. SANDISON, R. A., SPENCER, A. M., & WHITELAW, J. D. A. The therapeutic value of lysergic acid diethylamide in mental illness. *J. Ment. Sci.*, 1954, **100**, 491-507.
16. SLOANE, B., & DOUST, J. W. Psychophysiological investigations in experimental psychoses: results of the exhibition of D-lysergic acid diethylamide to psychiatric patients. *J. Ment. Sci.*, 1954, **100**, 129-142.
17. STOLL, W. A. Lysergsäure-diäthylamid, ein Phantasticum aus der Mutterkorngruppe. *Schweiz. Arch. Neurol. & Psychiat.*, 1947, **60**, 279-323.
18. STOLL, W. A., & HOFMANN, A. Partialsynthese von Alkaloiden von Typus des Ergobasins. *Helv. Chim. Acta*, 1943, **26**, 944-965.

133 East 58th Street  
New York City 22

## TRANSMISSION OF INFORMATION ABOUT SUCROSE AND SALINE SOLUTIONS THROUGH THE SENSE OF TASTE\*

*Department of Psychology, Harvard University*

J. G. BEEBE-CENTER, M. S. ROGERS, AND D. N. O'CONNELL<sup>1</sup>

### A. THE PROBLEM

Garner and Hake (2) have shown that Shannon's (4) measure of information transmitted through a channel of communication can readily be adapted to the measurement of information transmitted in absolute judgments. They use the term  $I_t$  to identify the measure so used. The present paper applies  $I_t$  values, corrected for bias,<sup>2</sup> to the measurement of information transmitted in absolute judgments on sucrose and saline solutions.

Three experiments were carried out. The first two concerned the transmission of information about *simple* solutions (i.e., solutions of sucrose or of salt); the third concerned transmission of information about *compound* solutions (i.e., solutions of sucrose and salt).

The procedure was essentially the same in all the experiments. The stimuli always covered a range of 100 gusts (1) in equal log gust steps.<sup>3</sup> The number of alternative stimuli in any one experimental series was known to the subject. The number varied from 3 to 17 in the different experiments. The absolute judgments involved identification of stimuli by number. After each identifying response *E* spoke the correct number and added "correct" if the response was correct. Stimuli were presented singly at approximately 2-minute intervals. *S* sipped some 5 to 10 cc of the solution from a small glass, expectorated, gave an identifying response, and rinsed his mouth.

\*Received in the Editorial Office on September 27, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This research was in part supported by contract N5ori-76, Project Order II, between Harvard University and the Office of Naval Research (Project NR142-201, Report PNR144). Reproduction for any purpose of the U. S. Government is permitted. W. Atkinson, E. Brush, T. Downer, R. Herrnstein, T. Lohr, W. Morse, E. Segal, and D. Zimmerman cooperated in this research.

<sup>2</sup> $I_t$  values computed from any sample are biased (3). We have corrected for bias either by the procedure developed by Miller and Madow (3) or by a procedure developed by one of us (Rogers) and as yet unpublished.

<sup>3</sup>Gusts are related in approximately linear fashion to concentrations. Thus the range corresponds approximately to 2 log units in terms of concentrations, and the steps to approximately equal spacings in terms of *jnd's*.



## B. EXPERIMENTS 1 AND 2: SIMPLE SOLUTIONS

A first experiment measured  $I_t$  for a single observer (B-C) for nine sucrose concentrations ranging from 1 to 115 gm sucrose per 100 cc tap water in  $\frac{1}{4}$  log gust steps. On each of 10 experimental days the nine stimuli were presented in three successive nonrecurrent chance series.  $S$  knew the schedule, though not, naturally, the chance series. In this experiment  $I_t$  was found to be 1.6880 bits. This means that  $S$  (B-C in this case) could differentiate with complete accuracy only a little more than three alternative sucrose concentrations.

A second experiment measured  $I_t$  for three, five, nine, and 17 concentrations of saline solutions ranging from 0.30 to 34.7 gm NaCl per 100 cc tap water in equal log gust steps. Two  $S$ s were used for each set of concentrations (see Figure 1). When nine saline concentrations were presented, the procedure was exactly as described above for the nine sucrose concentrations. For 17 and five concentrations the procedure was similar except that for 17 concentrations there were two nonrecurrent chance series on each of the 10 experimental days, and for five concentrations there were three chance series, each with single recurrence. The experiment with three concentrations was terminated after two experimental days, as neither subject had made a single mistake. On each day there had been 10 nonrecurrent chance series of three concentrations.

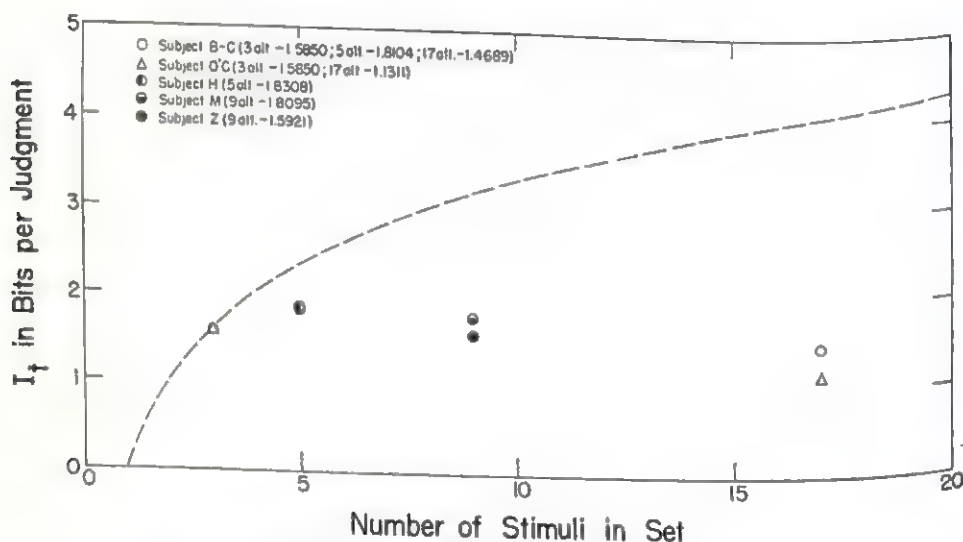


FIGURE 1  
INFORMATION TRANSMITTED IN ABSOLUTE JUDGMENTS OF TASTE INTENSITY AS A  
FUNCTION OF DENSITY OF STIMULUS DISTRIBUTION

The results of the second experiment are presented in Figure 1. The broken line represents values of  $I_t$  for perfect transmission ( $\log_2$  of the number of alternatives). It is quite clear that beyond five alternatives  $I_t$  does not increase with number of alternatives. Indeed,  $I_t$  appears to go through a maximum around four alternatives.

The values of  $I_t$  for nine saline alternatives were 1.8095 and 1.5921, as compared with 1.6880 for nine sucrose alternatives. It seems then that the same psychological range (100 gusts) of saline and sucrose solutions presented in the same number of equal log gust steps gives practically the same  $I_t$ .

### C. EXPERIMENT 3: COMPOUND SOLUTIONS

Sixteen compound solutions were prepared, representing all possible pairings of concentrations of 1, 4, 19, and 115 gm sucrose per 100 cc tap water with 0.30, 1.0, 4.8, and 34.7 gm NaCl per 100 cc tap water. Compounding involved addition of solutes, but not of solvents. Thus the compound representing 1 gm sucrose per 100 cc tap water and 0.30 gm NaCl per 100 cc tap water was made up by combining 1 gm sucrose, 0.30 gm NaCl, and 100 cc water. The concentrations of sucrose and salt involved represented ranges of 100 gusts divided into equal log gust steps.

Each of the 16 solutions was presented to each of five Ss twice, in chance order, on each of 15 days. The Ss knew that the solutions involved one of four saline concentrations combined with one of four sucrose concentrations. They were asked to name the solutions by identifying by number 1 (low) to 4 (high) first the sucrose concentration, then the salt concentration. The results are presented in Table 1. It is quite clear that  $I_t$  for identification of the 16 compound stimuli (i.e., computed from a 16 x 16 table) is very much the same as the sum of  $I_t$  for identification of the sucrose concentrations only (4 x 4 table) and  $I_t$  for identification of the saline concentrations only (4 x 4 table). Consequently,  $I_t$  is approximately summative in this case.

For all five Ss, however,  $I_t$  for the compounds exceeds slightly the sum of  $I_t$  for the sucrose concentrations plus  $I_t$  for the saline solutions. This would presumably happen by chance only once in 32 times. This needs further investigation.

Another question needing investigation concerns masking. In the experiment on simple saline solutions involving five alternatives the  $I_t$  values were 1.8104 and 1.8308. In the experiment on compound solutions, the  $I_t$  values for saline alone with four alternatives varied from 0.9780 to 1.5915 (see Table 1). There is no overlap in the two distributions. To what extent

TABLE 1  
 $I_t$  FOR COMPOUND SOLUTIONS OF SUCROSE AND SALT, AND FOR THEIR COMPONENTS

Subject	Sucrose alone	Salt alone	Sum	Compound solution
A	.8279	.9780	1.8059	1.8882
B	1.0754	1.0140	2.0894	2.2810
D	.6968	1.0128	1.7096	1.8039
L	1.3807	1.5915	2.9722	3.0173
S	1.0426	1.1163	2.1589	2.2598

is this difference due to masking, to what extent to the more complicated task of identifying the compound stimuli? A first step towards answering this question, it seems to us, is to get quantitative data on the masking interaction of supra-liminal concentrations of sucrose and salt. We hope to do so shortly.

#### D. SUMMARY

Transmission of information through the sense of taste about solutions of sucrose and salt was measured by means of Garner and Hake's statistic  $I_t$ , corrected for bias.

For simple saline solutions  $I_t$  does not increase as the number of alternatives is increased beyond five. It appears to go through a maximum around four alternatives.

$I_t$  for nine simple sucrose solutions is very nearly the same as for nine simple saline solutions.

For compound solutions of sucrose and salt  $I_t$  is slightly greater than the sum of the  $I_t$  values for the components computed from the same data.  $I_t$  values for the saline component alone are markedly less, however, than  $I_t$  values for five simple saline concentrations.

#### REFERENCES

1. BEEBE-CENTER, J. G. Standards for use of the gust scale. *J. of Psychol.*, 1949, 28, 411-419.
2. GARNER, W. R., & HAKE, H. W. The amount of information in absolute judgments. *Psychol. Rev.*, 1951, 58, 446-459.
3. MILLER, G. A., & MADOW, W. G. Information measurement for the multinomial distribution. In *Quarterly Progress Report*, H.R.R.L., Psychological Research Program, M.I.T., 16 December, 1952.
4. SHANNON, C. E. The mathematical theory of communication. In C. E. Shannon & W. Weaver, *The Mathematical Theory of Communication*. Urbana, Ill.: Univ. Illinois Press, 1949. (Pp. 3-91.)

Department of Psychology  
 Memorial Hall  
 Harvard University  
 Cambridge 38, Massachusetts

## A FACTORIAL STUDY OF THE COMPONENTS OF MUSIC\*

*Department of Music, University of California at Los Angeles*

---

ROBERT I. HENKIN<sup>1</sup>

---

### A. INTRODUCTION

The interest of the present study is in the isolation of musical factors in esthetic appreciation. The approach used for the investigation is a synthesis of previous techniques used in the factor analysis of music and esthetics. The importance of the work itself lies in the uncovering of musical factors which can be used as basic concepts for a general study of musical experience. The results of the analysis can be conceived as a set of factors which define a musical sound within an esthetic framework. The factors derived are in a general sense analogous to the physical correlates of sound (e.g., frequency, amplitude) or the subjective correlates of sound (e.g., pitch, loudness).

In a survey of the previous literature, one finds two main streams of thought concerning the definition of musical factors. One proposes an introspective analysis of the relationship between categories of affective experience and elements of musical form. The other attempts an analysis of factors by statistical processing of tests of musical aptitude or musical stimuli based on the measurement of thresholds. Studies which clearly demonstrate the use of factor analysis in the study of esthetics of music have not been found, but an understanding of the problems involved in the application of the method of factor analysis to the definition of esthetic factors of music can be achieved by considering studies which deal with factor analysis as applied to esthetic problems in general.

One of the more meaningful works utilizing introspective analysis and possessing some control of the stimulus material was written by Kate Hevner (8) in 1934. She reports an experiment in which "the affective value of certain elements of musical form has been studied by preparing two versions of a musical composition which differ in one respect only (e.g., rhythm, harmony, or melody) and presenting the two versions to two different groups of

---

\*Accepted for publication by Roy M. Dorcus of the Editorial Board, received in the Editorial Office on October 4, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>For their assistance during the preparation of this work the author wishes to thank Dr. L. A. Petran and Dr. J. N. Vincent of the department of music, and Dr. George E. Mount of the department of psychology of the University of California at Los Angeles, without whose help and support this work could not have been completed in its present form.



listeners who report their reactions by checking adjectives from a printed list. She went on to illustrate how certain musical ideas such as "firm rhythms . . . simple consonant harmonies . . .," brought about a characteristic selection of adjectives that appeared to be related to the ideas represented in the music.

Variations of this method appear in numerous publications, most characteristically in the work of Ralph H. Gundlach (6) in 1934 and Ivy G. Campbell (3) in 1942. Gundlach tried to obtain a consistent trend in statements of mood evoked by descriptive mood adjectives. Similarly Campbell tried to discover what emotions were evoked by playing musical selections and asking observers to designate the mood of the selection from a previously prepared list of adjectives.

In general, the conclusions of the above works point to the degree of correspondence between the lists prepared by the experimenter as representative of mood categories expressed in the music and what the subjects found as represented therein. Further they stressed the relative agreement of observers within the main categories of adjective descriptions and tried to explain the possible causes of any differences.

These studies give a fair cross-section of the introspective technique used to define music by the verbal communication of mood. The results obtained by this procedure are, however, of limited value due to the flaws which are characteristic of the method. Specifically, the accepted verbal patterns of response to given musical selections are not taken into consideration in the interpretations of the experimenter. The subject learns, as a part of his cultural tools, that certain words are accepted as characterizing certain musical ideas. Thus the minor keys are accepted as expressing a sad mood, the major mode, a happier mood, firm rhythms a triumphal mood, etc. The previous experiments do not show any necessary relationship between these verbal mood categories and the actual esthetic judgments on the part of the subjects. The method requires only that the subject recognize and choose the accepted verbal characterizations of the music.

This technique is even used unwittingly at times to bias the subject during the instruction period by emphasizing the conventionalities of learning concerning musical expression. One must seriously ask if the agreement pointed out by the previous experimenters is not superficial, dealing only with agreement on a learned verbal level and not indicating the true nature of the subjects' experience. For, in reality, what meaning is there in the agreement of a subject group which is instructed to follow their learned responses, then hears a given selection and answers in the way in which they were originally biased.

Statistical factor analysis has been used in connection with tests of musical aptitude and psychophysical judgments of sound, but not in connection with the more general esthetic problem of music. The problems that have been encountered in the use of statistical factor analysis in other esthetic areas are similar to those encountered in the present study and thus will be considered also. The importance of each of these approaches, although not directly related to the present work, lies mainly in its application of method.

One of the first factor analyses of music appeared in 1939 when Raleigh M. Drake (2) factor analyzed eight different music tests, using the Spearman tetrad-difference technique. These tests included five from the Seashore battery measuring pitch, rhythm, intensity, time and tonal memory, respectively, one from the Kwalwasser-Dykema battery called a test of tonal movement, and two constructed by the author emphasizing musical memory and retentivity. Drake found one common factor and a suggestion as to its nature was stated.

In a similar fashion H. D. Wing (11) and J. R. Karlin (10), both in 1941, published reports of factor analyzing other music tests by which they could define some general factor of music ability. Karlin, in addition to this, was interested in finding a general auditory factor (9).

Wing's work is somewhat obscure for he has neither clearly stated the form of his stimulus material nor the criteria used in scoring and it is thus difficult to know the form and meaning of the factors which he derived. It is possible that he was measuring discrimination factors similar to those involved in other musical analyses or that some esthetic factors were involved. The former is perhaps the better possibility since the intercorrelations with clearly discrimination test materials, which were also included, were significant and appeared in the same derived factors.

Karlin followed the same general pattern of procedure as outlined by Drake above. He achieved no general music ability factor but three independent discrimination factors of tonal sensitivity, memory for musical elements, and memory for musical form. He further factor analyzed tests of auditory function in the fields of pitch, loudness, timbre, etc., looking for a general factor of audition. Here he obtained eight group factors including pitch-quality discrimination, loudness discrimination, and memory span.

One of the first studies using factor analysis in the general area of esthetic appreciation was published in 1940 when H. J. Eysenck (4) stated that he had conclusively found "a general, objective factor of visual aesthetic appreciation which is independent of teaching, tradition, and other irrelevant asso-

ciation." His experimental procedure consisted of presenting 18 sets of picture cards in decks of 15 to 18 diversified observers. These cards, which were mostly new esthetic experiences for his subjects, were drawn from portraits painted by the great classic masters, statuary of famous emperors, pictures of masks, vases, and similar materials. The subject were then to rank these cards according to the criterion of how well they liked each picture. The rankings were intercorrelated and then factor analyzed. From this he found a common factor which contained most of his experimental variance and which he labeled "a general objective factor of aesthetic appreciation." The general procedure outlined in this experiment uses meaningful choices in preference to verbalizations lacking behavioral reference. This basis concept, coupled with the improvements in technique found in later works, provided the major part of the methodological basis for the present study.

Guilford and Holley (5), in 1949, used a more sophisticated approach than in the past, although using the same general technique as illustrated above. Their greater care and control led to the extraction of more clear-cut factors which could be more easily interpreted. Working with the esthetic judgments made in reference to the designs on the backs of playing cards, Guilford applied certain standard procedural techniques to esthetic materials which previously had not been used either effectively or consistently: (a) he used the method of equal appearing intervals, (b) he set up artistic criteria upon which the observers based their judgments, and (c) he tried to interpret the results in terms of esthetic factors which stem, in general, from the observations of the designs on the backs of the playing cards themselves.

It is thus apparent that although there have been related studies there has not yet been a serious attempt made at a psychological interpretation of music from an esthetic reference using statistical methods. It is hoped that this paper will combine the work done in the areas previously mentioned in order to develop independent factors which define the elements constituting the music piece itself.

## B. THE EXPERIMENT

### 1. *Material and Procedure*

Factor analysis involves the extraction of clear-cut factors in the responses made to a defined situation. This requires that the situation be capable of such organization by the subject. It further requires that the experimenter be capable of interpreting and giving meaning to the particular organization of the data found in the factor analysis. Most factor analysts have approached

this problem by choosing the stimulus material used in a specified manner or form and then using the factor analysis method to find the degree to which the factors located in the subject responses are of the same form. This procedure is proper provided the experimenter does not bias the subject response in such a manner as to determine the form of the results obtained (i.e., the choice of stimuli and subject response are independent), and provided he can show that the result is repeatable.

This was the general procedure used in the present study. The stimulus pieces were explicitly chosen to provide a clear-cut basis for organization by the subject if such could be found to exist. This meant selecting records which exhibited, if possible, one factor per type of music piece. For each of three factors, three music pieces loaded with each of the supposed factors were used to represent that factor in the experimental procedure. This was done in order to yield a general definition of each factor. In order to emphasize the timbre differences among the records, a tenth record was so inserted that if any variance dependent upon orchestral color were isolated it would correlate in a large degree with that of the added recording. Thus the function of the factor analysis was to determine whether and to what extent this predicted form was the correct one.

One will see in the procedural part of the experiment that the musical factors were obtained from purely affective responses on the part of the listeners, i.e., the subjects were instructed to state a degree of preference rather than to choose from among a list of adjectives or otherwise to attempt a verbal description. There was neither biasing of the subject choices by using adjectives describing preferences nor was there any other form of instruction used which would determine the subjects' choices. It is thus assumed that from the measure of how well a subject liked a given music piece we can define the major esthetic elements that make up the music.

It is perhaps important to emphasize that the musician's viewpoint of music as a configurational whole is not violated by this approach for its purpose is not to destroy the effect of music as a gestalt by breaking it down into isolated elements; rather it is an attempt to understand the manner in which esthetic experience is determined in an actual music setting. When we use the nomenclature "classical" and "modern" or "jazz" and "serious" music we assume some similarity or difference which can be shown to exist because we can observe it in the music. This makes a classification by which we can better identify diversified music pieces. In a like fashion this study attempts to discover the similarities in a group of music pieces on the basis of listener preference and tries to classify all music in the light of



this systemization. The value of such a classification lies in how great a generalization about music can be made using it as a basis.

In the experiment itself there was (a) a set of pre-test or practice trial records and (b) a set of experimental or complete presentation records. The records in both groups were picked on the same basis, that they exhibit a preponderance of one factor over others which will make its presence more easily definable than if all elements were equally weighted. Each set of records was played twice, separated by a period of four weeks, to approximately the same subject group. The records used in the two practice trials differed from those used in the complete presentation and from each other with one exception, to be described later. The records of the complete presentation were the same at both trials but presented in opposite order. On Trial I they were used in order 1-10, on Trial II, using the same numbering system, in order 10-1.

The records were chosen, wherever possible, in reference to the time of composition in the history of music. A cross-section of musical taste from the three periods of (a) Baroque, pre-classic, and classic, (b) romantic, and (c) modern was attempted. This was done to avoid needless duplication in affective responses within any one time category familiar to the average person and to increase the generality of the results.

The order in which the records were played was partially systematized. In the practice trials as great a difference between records as possible was attempted in order to prepare the listener, with respect to style and period of composition, for the coming series of differences in the regular experiment. The complete presentation was so arranged that as far as possible there was a continual shifting of emphasis between guessed factors present and time period of the piece, so that neither pieces from the same period of composition nor pieces containing similar stressed elements were placed in succession.

The records of practice Trial I are as follows:

1. *Khatchaturian, A.* Gayne Ballet Suite—Dance of Ayshe; Columbia mm. 664 (78 rpm). Played by Philharmonic Symphony Orchestra of New York, conducted by Efrem Kurtz.

This record was chosen as a representative of the romantic-modern period with the emphasis placed on melody. Its task as the first record in the experiment was to settle the auditors into the coming procedure by making them feel at home musically in a very familiar, semi-popular style.

2. *Vivaldi, A.* Concerto in G Major, "alla Rustica," for strings and cembalo—2nd movement; Decca DL9575 (33 1/3 rpm). Played by the Virtuosi di Roma, conducted by Renato Fasano.

This record was chosen as a representative of the Baroque period with the emphasis on harmony.

3. *Varèse, E.* Ionization—approximately one-half of the composition was played; E.M.S. 401 (33 1/3 rpm). Performed by the Juilliard Percussion Orchestra with Frederic Waldman conducting.

This record was used primarily as an example of rhythm and it was written in the modern style of composition. It was "different" enough from the usual run of music pieces to give the listeners some basis for judging the remaining records in the complete presentation, by giving them a reference point in affective response.

The records of practice Trial II are as follows:

1. *Khachaturian, A.* Gayne Ballet Suite—Dance of the Rose Maidens; Columbia mm. 664 (78 rpm). Played by Philharmonic Symphony Orchestra of New York, conducted by Efrem Kurtz.

This record was intended as an alternate form of the corresponding record in practice Trial I. It was loaded, factorially speaking, in the melodic factor and from the point of time development lay in the romantic-modern period.

2. *Honegger, A.* Pacific 231 (one side of a 10" record); Victor 276- (78 rpm). Performed by Continental Symphony Orchestra, Pierro Coppola conducting.

This record was used as an alternate form of record No. 3 on practice Trial I. It was primarily loaded in rhythm and belonged to the modern period.

3. *Vivaldi, A.* Concerto in G Major, "alla Rustica," for strings and and cembalo—2nd movement; Decca DL 9575 (33 1/3 rpm). Played by the Virtuosi di Roma, conducted by Renato Fasano.

This record was the same as that used in practice Trial I. It was used twice since no acceptable alternate form of this record within this style period could be easily found. Again, it was loaded harmonically and belonged to the Baroque period.

The records of the complete presentation are as follows:

The records of the complete presentation are as follows:  
1. *Schubert, F.* Symphony No. 2 in B Flat Major—2nd movement (full statement of the theme and one complete variation); Columbia M 420 (78 rpm). Performed by Columbia Broadcasting Symphony, Howard Barlow conducting.

This record was chosen as a representative of the classic period, primarily loaded in melody. Its secondary loading was hypothesized to be harmony.

2. *Stravinsky, I.* The Rite of Spring—Sacrificial Dance; Columbia

LP ML 4092 (33 1/3 rpm). Played by Philharmonic Symphony Orchestra of New York, Igor Stravinsky conducting.

This record was used to show loadings in rhythm as an example of the modern period. There was a small secondary loading in melody assumed also.

3. *Griffes, C.* The White Peacock (one side of a 10" record); Columbia 17140-D (78 rpm). Performed by Howard Barlow conducting Columbia Broadcasting Symphony.

This record was chosen to show an orchestral color factor. Since all but one of the pieces chosen were orchestral works the presence of a large amount of color factor in this selection was hypothesized to correlate with certain other of the pieces and thus define an orchestral color element. The piece, written in a romantic-modern style, is assumed to have secondary loadings in both harmony and melody.

4. *Mendelssohn, F.* Symphony No. 4 in A Major—4th movement (one side of a 12" record); Victor M 119 (78 rpm). Performed by La Scala Orchestra, Milan, Ettore Panizza, conductor.

This record was chosen to show a predominantly rhythmic piece from the romantic period.

5. *Vivaldi, A.* Concerto in A Major, for strings and cembalo—2nd movement; Decca DL 9575 (33 1/3 rpm)—Performed by Virtuosi di Roma, Renato Fasano, director.

This record was chosen to represent harmony. It was written in the Baroque style. Some melodic elements were also assumed.

6. *Debussy, C.* Nocturnes—Nuages (one side of a 12" record); Columbia M—344 (78 rpm). Played by D. E. Inghelbrecht conducting Grand Orchestra Des Festivals Debussy.

This record was chosen as being predominantly loaded in harmony. It was written in the romantic-modern style and it was assumed to have secondary loadings in melody.

7. *Mossolov, A.* The Steel Foundry (one side of a 10" record); Columbia M—347 (78 rpm). Performed by Orchestra Symphonique of Paris, Julien Ehrlich, conducting.

This record was predicted to be loaded primarily in rhythm. It represents a modern style of composition.

8. *Bach, J. S.* Prelude No. 1 in C Major from the W.T.C., Vol. I; Victor M—1106 (33 1/3 rpm). Performed by Wanda Landowska, harpsichordist.

This record, the only one which was not an orchestral work—played on

the harpsichord—was thought to contain mostly a harmonic element. It lay within the Baroque period from the point of view of musical time development.

9. *Kabalevsky, D.* Symphony No. 2 in C minor—2nd movement (from the beginning up to the end of the exposition section); Capital Classics L 8032 (33 1/3 rpm). Performed by Jacques Rachmilovich conducting the Symphony Orchestra of the Academy of Santa Cecilia, Rome.

This record was used as a representative of the melodic factor assumed, chosen from the modern period. It was hypothesized to be loaded slightly in harmony also.

10. *Tschaikowsky, P.* Symphony No. 4 in F minor—2nd movement (one side of a 12" record); Victor 327 (78 rpm). Performed by Serge Koussevitsky conducting the Boston Symphony Orchestra.

This record was chosen to represent the melodic factor in the romantic era of composition.

The records, systematized by predicted primary factor loading and period of composition, appear in Table 1.

PREDICTED PRIMARY FACTOR				
Period	Rhythm	Melody	Harmony	Color
Classic or Pre-Classic	No records available	Schubert 2nd movement 2nd symphony	Vivaldi 2nd movement Concerto in A major Bach C Major Prelude of Well Tem- pered Clavier	
Romantic	Mendelssohn 4th movement 4th symphony	Tschaikowsky 2nd movement 4th symphony	No records available	Griffes The White Peacock
Modern	Stravinsky Rite of Spring	Kabalevsky 2nd movement 2nd symphony	Debussy Nuages	

A general type of rating sheet containing a key or scoring formula, a practice trial rating schedule, and a complete presentation rating scheme was distributed to 243 various majored students enrolled in a music appreciation course at U.C.L.A. General instructions were given to acquaint the students with the three parts of the test sheet, defining the meaning of the terms pleasant and unpleasant and stating major characteristics of the form of relative scale used.



The instructions further outlined the whole procedure of the specific experiment indicating the number of records to be played [13] and further emphasizing the relative characteristics of the scoring formula.

The practice trial was run to acquaint the subjects with the use of the scoring formula and the judgment procedure in general. After answering questions the experimental records were presented.

After all the records were played some measure of each subject's familiarity with the records was obtained by direct questioning as to the number of records the subject had heard before.

Four weeks later the same procedure was applied to the same class of students. This time only 193 students participated in the work. The population was essentially the same, however, although absences accounted for some difference in the number of students who took the second examination. The work was carried on in the same classroom, at the same time, and under as similar conditions as possible to those in the first experiment. The instructions read approximated those read at the first presentation.

After the last record of this trial was played a biographical data sheet was distributed to the subjects, carefully explained by the experimenter, and filled out by the subjects.

## 2. *The Statistical Procedures*

Pearson product moment coefficients were obtained for each part of the response ratings on each administration of the test records.

Table 2 shows correlation Matrix I obtained from the first administration of the tests.

TABLE 2  
CORRELATION MATRIX I OBTAINED FROM THE FIRST ADMINISTRATION OF THE TESTS

	1	2	3	4	5	6	7	8	9	10
1	1.000	-.045	.076	.383	.277	.223	-.108	.202	.131	.414
2		1.000	.205	.002	.120	.181	.534	-.048	.082	-.028
3			1.000	.153	.045	.345	.077	.088	.408	.234
4				1.000	.240	.114	-.028	.136	.199	.443
5					1.000	.297	.172	.199	.124	.349
6						1.000	.113	.104	.457	.271
7							1.000	.017	-.012	-.098
8								1.000	.133	.247
9									1.000	.418
10										1.000

Table 3 shows correlation Matrix II obtained from the second administration of the tests.

The resulting correlation matrices were processed by means of a variation

of the principal components method of extraction of factors and an attempt was made to verbally relate the factors arising from this analysis to the form of the music chosen for the study.

TABLE 3  
CORRELATION MATRIX II OBTAINED FROM THE SECOND ADMINISTRATION OF THE TESTS

	1	2	3	4	5	6	7	8	9	10
1	1.000	-.319	.066	.302	.239	.022	-.330	.200	-.145	.211
2		1.000	.204	-.153	.008	.282	.698	-.066	.278	.012
3			1.000	-.012	.015	.394	.216	-.077	.411	.297
4				1.000	.221	-.005	-.177	.116	-.052	.070
5					1.000	.088	-.099	.314	.100	.085
6						1.000	.303	.045	.420	.256
7							1.000	-.107	.209	.062
8								1.000	-.091	-.080
9									1.000	.281
10										1.000

The factor analysis variation used was one adapted for use on I.B.M. equipment and involved the determination of eigenvalues of symmetric matrices. The method was originated by Hestenes and Karush in 1951 (7) and brought into practical application in the paper of Everett Yowell (12) also in 1951.<sup>2</sup>

The method of extraction relieved us of the responsibility of rotation after extraction since an optimum solution had been reached by solving for the eigenvalues. Any rotation would have consisted mainly in a juggling of the variances of the given factors and would have lent nothing toward the ultimate goal of rotation, i.e., positive manifold and simple structure.

The factor matrix obtained from correlation Matrix I appears in Table 4.

The factor matrix obtained from correlation Matrix II appears in Table 5.

The purpose of repeating the experiment twice can now be more clearly seen since we have an indication as to how reliable the test records really were. In general one can see that there are two factors clearly defined, a third fairly well defined, and three residual factors. Further, the consideration of variance can be shown to be important since it underlines similarities in corresponding places on both factor matrices. One can note that Factor A in Matrix F<sub>I</sub> corresponds to Factor B in F<sub>II</sub> and that in a like manner, Factor A in Matrix F<sub>II</sub> corresponds to Factor B in F<sub>I</sub>. This can be clearly seen by a comparison of the factor loadings in each of the individual tests comprising each factor. This reversal of variance deserves close at-

<sup>2</sup>The author is indebted to Owen Mock for his assistance in the statistical processing of the data.

tention since it is indicative of what sophistication on the part of the auditors may produce in the way of sharpening of affective responses. When first in contact with an unfamiliar music piece the listener appears to place all his attention on the melodic content and responds in an affective manner

TABLE 4  
FACTOR MATRIX,  $F_I$  OBTAINED FROM CORRELATION MATRIX I

	$A_I$	$B_I$	$C_I$	$D_I$	$E_I$	$F_I$
1.	.34	-.27	-.27	-.20	-.05	-.72
2.	.11	.61	-.18	-.17	.18	-.14
3.	.30	.23	.46	-.03	.33	-.17
4.	.35	-.20	-.21	-.47	.40	.21
5.	.33	.05	-.45	.12	-.49	.33
6.	.38	.21	.23	.12	-.51	-.27
7.	.05	.59	-.35	.02	.37	.08
8.	.23	-.13	-.21	.82	.41	-.08
9.	.39	.08	.47	.06	-.04	.29
10.	.45	-.21	-.03	.09	.07	.31
V.	2.76	1.72	1.24	.91	.82	.66

TABLE 5  
FACTOR MATRIX,  $F_{II}$  OBTAINED FROM CORRELATION MATRIX II

	$A_{II}$	$B_{II}$	$C_{II}$	$D_{II}$	$E_{II}$	$F_{II}$
1.	-.24	.47	-.09	-.12	.46	.01
2.	.46	-.14	-.36	-.27	.10	-.14
3.	.35	.29	.24	.10	.11	-.37
4.	-.17	.36	-.10	-.75	-.28	-.32
5.	-.05	.41	-.48	.08	-.39	.44
6.	.38	.30	-.04	.13	.13	-.34
7.	.46	-.18	-.27	-.31	.28	.11
8.	-.14	.25	-.59	.39	.30	-.18
9.	.40	.23	.12	.25	.55	-.05
10.	.19	.38	.38	-.06	.22	.62
V.	2.59	1.92	1.25	.86	.77	.72

to the presence or absence of this melodic invention. However, this shift in variance illustrates the possibility that on successive listenings, the auditor bases his affective response not solely on the melodic characteristics of the selection, but on other elements of musical composition, some of which are enumerated later. This sharpening of response can be seen again in the subject response to a smaller number of music pieces in a more characteristic manner. This takes the form, as seen in most corresponding factors, of a listener reacting in a stronger manner, either positively or negatively to a piece which is more familiar to him, than to one which is unfamiliar to him.

### 3. *A Verbal Interpretation of the Factors*

The elements of melody, harmony, rhythm, and orchestral color have been mentioned up to now only in passing, and deserve a more careful analysis and definition that they may have more meaning when applied as generic names of musical factors. The definitions which follow stem primarily from a musical interpretation of the factors and are limited in their psychological application by this fact.

*a. Harmony.* Any simultaneous combination of musical sound. This could be implied, as in arpeggios which are harmonies spread out in time, or realized, as in chords.

*b. Melody.* A series of tones rising and falling by definite pitch intervals. It is in this sense very difficult to separate melody from rhythm verbally since musical sound has two definite qualities, pitch and duration, both of which enter into the making of the melodic line. Hence this rising and falling of a series of tones is built upon a rhythmic background marked by a certain stress or accent laid upon certain tones at organized time intervals. Melody is essentially rhythm and variation in pitch.

*c. Rhythm.* Organization of movement (in this case movement of musical sounds) into measureable values. In reality a definition of rhythm in the verbal sense is nearly impossible since the term is so general one cannot pin it down to a specific task in music.

*d. Timbre* (orchestral color). The quality or "color" of a tone, i.e., the difference between tones of the same pitch if produced on various instruments, e.g., a flute and a violin (1). In this work the term "color" was applied to an orchestral sound in general which tended to produce an affective state in the listener which would not have occurred if the sound were produced on a piano or some similar single instrument, without the complex overtone patterns of a symphony orchestra.

To draw a more careful picture of the factors involved, let us look at what factors are indicated by the previously mentioned numbers.

### C. ANALYSIS OF THE FACTOR MATRICES

#### 1. *Factor A*

In Factor *A* we have significant loadings in Matrix  $F_1$  in Tests 1 (the 2nd movement of Schubert's *Symphony No. 2*), 3 (Griffes' *The White Peacock*), 4 (the 4th movement of Mendelssohn's *Symphony No. 4*), 5 (the 2nd movement from Vivaldi's *Concerto in A Major* for violin and orchestra), 9 (the 2nd movement of Kabalevsky's *C Minor Symphony*), and 19 (the 2nd movement of Tschaikowsky's *Symphony No. 4*). This appears to



be the main factor in Matrix  $F_I$  contributing 27.6 per cent of the total common factor variance. In the remaining tests, there appear to be no appreciable loadings in the factor under consideration except in Test 8, Bach's *Prelude No. 1* from the W.T.C., Vol. 1, which has a slightly significant loading.<sup>3</sup>

In the corresponding factor, Factor  $B$ , in Matrix  $F_{II}$  there are significant loadings in Tests 1, 3, 4, 5, 6, and 10 (all corresponding to the above). There are also slightly significant loadings in Tests 8 and 9. This is a secondary factor in Matrix  $F_{II}$  contributing 19.2 per cent of the total variance. One can easily see the high degree of correspondence between these factors in the matrices, the only difference between the two administrations appearing to be in a small change of the degree of strength of the factor loadings of some corresponding factors.<sup>4</sup>

The remaining records appear to have no function in the determination of this factor and because of the rather clear-cut delineation in the records I chose to call this primary factor a *Melody Factor*, since the records which came out with the highest loadings were those previously in the melodic factor.

## 2. Factor B

In Factor  $B$ , Matrix  $F_I$ , we find very significant factor loadings in Tests 2 (Stravinsky's *Rite of Spring*) and 7 (Mossolov's *The Steel Foundry*). There are slightly significant loadings in Tests 3 (Griffes' *The White Peacock*) and 6 (Debussy's *Nuages*). Meanwhile, there are slightly significant negative loadings in Tests 1 (2nd movement of Schubert's *Symphony No. 2*) and 10 (Tschaikowsky's *4th Symphony, 2nd Movement*). This factor, a secondary factor, contributes 17.2 per cent of the total common factor variance.

There is a similar situation in Matrix  $F_{II}$  in Factor  $A$  where we get very significant loadings in 2 and 7 and significant loadings in 3, 6 and 9. We also get a slightly significant negative loading in Test 1, as before. The other loadings appear slight enough to be called insignificant in interpreta-

<sup>3</sup>The term significant is interpreted to mean any factor loading between .3 and .4. Very significant is a term applied to any loading above .4. A factor loading of .2 is called slightly significant while any loading less than .2 is referred to as insignificant.

<sup>4</sup>This change in quantity of loading says nothing against the accuracy of either the method or the retest reliability of the 2nd matrix. It does indicate, however, that the measures obtained are only approximations of the population coefficients of correlation and their change from test to test is to be expected. However, a good indication of the reliability of the method does lie in the good approximations of the correlation coefficients which were achieved as shown by the high retest reliability of the 2nd matrix.

tion. This factor is the primary factor in  $F_{II}$  and contributes 25.9 per cent of the common factor variance.

This delineation seems to lend itself to the interpretation of this factor in the direction of a *Rhythm Factor*, since the records which were chosen as rhythmic records find themselves highly emphasized in this factor.<sup>5</sup>

One might look for an interpretation of this factor in view of "modern music" (Records 2, 3, 6, 7, and 9) as against "classic music" (Records 1 and 10). This could be a possibility, but in view of the only slight negative correlation between these two groups of records and the fact that not all the records which lie in the "classic vein" take part in this negative correlation, I am more inclined to look at the factor as one of rhythm. There appear to be too many inconsistencies to hold the aforementioned view, more than if the rhythmic interpretation is held.

However, the interpretation of this factor is not as clear-cut as that of Factor *A*.

### 3. Factor *C*

Factor *C* in Matrix  $F_I$  has a very significant loading in Test 3 (Griffes' *The White Peacock*) and a significant loading in Kabalevsky's *2nd Symphony, 2nd Movement*. It has very significant negative loadings in 5 (Vivaldi's *Violin Concerto in A Major*) and 7 (Mossolov's *The Steel Foundry*). There is a slightly significant positive loading in Test 6 (Debussy's *Nuages*) and a slightly significant negative loading in Test 1 (Schubert's *2nd Symphony, 2nd Movement*). The remainder of the tests have small enough loadings so that the amount of loading which they possess could have occurred by chance. This factor is a tertiary factor in  $F_I$  and contributes 12.4 per cent of the total variance.

In Matrix  $F_{II}$  in Factor *C* there is a significant loading only in Test 10 (Tchaikowsky's *4th Symphony, 2nd Movement*) with a slightly significant loading in Test 3. The negative loadings, in general similar to those in Matrix  $F_I$ , lie in Tests 5, 8, 2, and 7. This factor is, as in the previous matrix, the third factor in  $F_{II}$  and contributes 12.5 per cent of the total common factor variance.

From the characterization of the records lying within this factor one could

<sup>5</sup>The failure of Test 4, the Mendelssohn 4th movement to show itself in this factor probably stems from the fact that although this piece is rhythmically motivated, written in a quick running motion, it is characteristically a melodic piece, and if we can believe the delineation in this factor, there appears to be a negative correlation between melody and rhythm. This negative correlation between records which appeared as melodic in Factor *A* and rhythmic in Factor *B* seems to be a characteristic of this factor.

expect this to represent an *Orchestral Color Factor*. The key appears to be the high positive loadings in those records which contain a great deal of color instrumentation, e.g., Griffes and Kabalevsky, as against the negative loadings experienced in records which do not contain as complex overtone patterns, most importantly, Test 8, the harpsichord recording of the 1st prelude in C Major from Bach's W.T.C. Vol. I.<sup>6</sup>

#### 4. Factor D

This factor is characterized in  $F_I$  by having a very significant positive loading on Record 8 (Bach's *Prelude No. 1* from the W.T.C. Vol. I) and a very significant negative loading on Test 4, Mendelssohn's *4th Symphony, 4th Movement*). These two tests seem to define this residual factor very well since the remaining loadings are all insignificant.

In the Matrix  $F_{II}$  the factor has the same characteristic loadings, significant in a positive direction on Test 8 and significant negatively on Test 4. In this matrix there is also a significant negative loading in Test 7, Mosolov's *The Steel Foundry*. All the remaining factor loadings are insignificant. Here, as in the previous three factors, one can see the close correspondence between Matrices  $F_I$  and  $F_{II}$ .

An adequate verbal interpretation, despite the apparent consistency in the results of the two matrices, is somewhat nebulous. Since there is such a great consistency between test and retest, one cannot throw away this factor even though it appears to be a residual factor, contributing only 9.1 per cent and 8.6 per cent of the total variance, respectively. However, an accurate definition of this factor from a verbal point of view appears unlikely.

#### 5. Factor E

Factor *E* in Matrix  $F_I$  contributes only 8.2 per cent of the variance of the total, and in the same fashion as Factor *D*, is a residual factor. There are, however, significant positive loadings on Tests 3, 4, and 8 and very significant negative loadings on Tests 5 and 6. The other loadings are insignificant and could have occurred by chance.

Similarly Factor *E* in Matrix  $F_{II}$  is a residual factor contributing only 7.7 per cent of the total variance. It has significant positive loadings in Tests 1 and 9 and a significant negative loading in Test 5. Here one can

<sup>6</sup>The Stravinsky piece does not take too much exception to this interpretation since the portion of the record played was more rhythmically barbaric than beautiful in terms of orchestral color. It was hoped to be in a classification, factorially speaking, with Mosolov's *The Steel Foundry* and it was further hoped it would measure approximately the same factor.

see the faltering of the reliability of the two matrices, due primarily to the fact that the factor under consideration is a residual factor and does not contribute much in the way of total common factor variance.

## 6. Factor F

The sixth factor is a completely residual factor, extracted for what intelligence it might contribute to the whole and cannot receive a meaningful interpretation even on this level of explanation.

These factors, *E* and *F*, both appear to be genuine residual factors and cannot be interpreted with any reliability due to their primary function as such. They were included for information purposes, not for any value of definition which they might contain.

The extraction of factors was stopped at this point since mathematical consideration of the value of the remaining factors was so slight that their importance was negligible.

## D. INTERPRETATIONS

### 1. Assumptions and Limitations

It is important to realize that any interpretation is tentative and is based on assumptions which can be tested adequately only by further experimentation. Thus in interpreting the results one must be careful to weigh the limitations imposed upon the generality of the study by the nature of the experimental design. For this reason we shall now approach the previously presented problem to ascertain what assumptions were involved, primarily on the experimental level. These limitations are dependent upon specific conditions of the experimental situation which may or may not affect the generality of the results. The major parameters are listed as follows, and at the same time, the applied methods of control or the method of determining the extent of the limitation is indicated.

#### a. The musical stimulus.

- (1). *Method of producing the musical stimulus.* Limited, since only commercial recordings were used. It is possible that different results would have occurred if different recordings were used or if live music were utilized.
- (2). *Length of recordings.* Limited by the attention span of the subject group and by the necessity that the experiment be completed within the allotted class time of 50 minutes. Thus the average playing time was about  $3\frac{1}{2}$  minutes with the extremes represented by the Vivaldi pieces (2 minutes) and the Stravinsky piece (5 minutes).
- (3). *Quality of recordings.* Limited by the varying qualities of re-



cordings used since some records were made at 78 rpm, some microgrooves at 33 rpm. Thus there may have been variations in the quality of the recordings with respect to fidelity, surface noise, and similar reproducing problems which might have had some effect on the interpretations made. This does not seem probable, however, since the factor patterns do not follow any apparent variation in quality.

(4). *The specific selections played.* It is possible that the particular selections chosen for this experiment do not represent music in any general way as they are intended. This element was considered as well as it was thought possible in a single experiment. A check on the assumptions involved here would require additional investigation.

(5). *Timbre differences between orchestral and non-orchestral selection.* It was thought that differences between orchestral and non-orchestral pieces might have an effect upon the results obtained. A harpsichord selection was therefore inserted in order to emphasize the appearance of such differences and permit their direct observation if they existed. This procedure resulted in the isolation of an orchestral color factor.

#### *b. Familiarity.*

(1). *Familiarity of subjects with the selections played.* Limited, since the best evaluation of the results could be obtained if the recordings were new esthetic experiences for the subjects. In this way the subject group would respond with approximately uniform judgments, from an experiential point of view, and this would lend greater consistency to the responses. The effect of familiarity was tested by obtaining a measure of the subjects' familiarity with the recordings at each administration of the tests by direct questioning after each examination. After the first administration only 3.7 per cent of the subjects were familiar with over half of the records played. After the second administration 20.2 per cent stated a familiarity with over half of the records. This recall is perhaps somewhat less than expected since the complete presentation records were exactly the same recordings as those portions played at the previous session one month earlier. However, the increased familiarity coupled with the similarity of major factors found in the two administrations indicates an important degree of independence of the results with respect to subject familiarity with the music.

(2). *Familiarity of subjects with the general style of composition exhibited by the recordings.* Limited in generally the same way as in the previous category. The effect was tested by direct questioning of the subject group after the second administration of the tests. The results show

a very close similarity to the familiarity response requested above since 21.7 per cent of the subjects expressed a familiarity with the type of music played in over half of the recordings. Another indication of familiarity with musical style was attempted by questioning concerning formal musical instruction of the subjects. This inquiry showed a rather high degree of musical experience among the subject group with 43.5 per cent having received formal musical instruction in either an instrument, voice or theory for two years or over.

(3). *Bias due to prior experience with the range of music used; i.e., "jazz" or popular music as opposed to "classical" or serious music.* Limited since only one level of music was played. If all levels, jazz as well as serious music, were included, the task of rendering accurate personal judgments, on the part of the subjects, would be greatly limited by prejudice and by an indeterminable varying degree of musical experience. This problem was compensated for in part by including a semi-serious selection from Khachaturian's *Gayne Ballet Suite* during the formative practice trial period to set in motion this "popular" side of the subjects' rating scale, yet keeping within the accepted framework of the remainder of the experimental records.

c. *Population effects.*

(1). *Differences in affective response to the musical stimuli due to the differences in training, educational experience, and similar cultural diversities.* Limited since familiarity was wished to be kept at a constant minimum level. These differences, although recognized, were not compensated since only college students were used as the subject group.

d. *Social situation.*

(1). *Differences between group listening or individual listening.* Limited since the effects of this social situation were not known. However, due to the nature of the experiments only group listening was possible and thus no indication of any differences could be uncovered.

(2). *Interaction between members of the subject group with respect to their responses to the musical stimuli.* Limited since the expression of each individual was necessary without communication between the members of the subject group. This was important since any uncontrolled or spurious interaction would lead to unreliability of the results. Interaction was kept to a minimum by careful supervision of subjects by proctors and by stressing this point in the instructions read to the subject group at the beginning of each experimental session.

*e. Environmental variables.*

(1). *Effects of temperature, room conditions, time of experiment, etc., upon the form of the results.* Since unsuspected forms of error may bias the judgments of the subjects these variables of temperature, room conditions, instruction forms, general procedure, etc., were controlled during the two administrations of the tests by running the experiment under as similar conditions as possible. The work was thus carried out in the same room, at the same time, etc., with as similar a procedure as possible.

## E. CONCLUSIONS

The factor solution that was obtained pointed to two definite conclusions:

1. That there is clear-cut evidence for two independent, meaningful factors that have been extracted from the aforementioned series of music pieces. These have been interpreted as (*a*) a melodic factor, and (*b*) a rhythmic factor. There is the possibility of a third, which could be interpreted as an orchestral color factor.

2. The music pieces chosen permit a tentative definition of the factors isolated in the factor analysis. Since single pieces did not appear, however, as pure measures of any of the factors the definitions are, to this degree, limited. Such a result is almost invariably the case and emphasizes the value of the factor analysis procedure since it helps clarify our initial attempts at analysis. In retrospect, we must say that we have established the existence of a number of independent, meaningful, musical factors. We must now proceed to a better approximation and definition of the factors by additional choice of stimulus tests and study of the resultant form of the data.<sup>7</sup> To do this in the simplest way would be to take the factors which are not well defined and using the records which possessed the highest factor loadings in this

---

<sup>7</sup>One notices that although a harmony factor was predicted, none was evident in the interpretation of the factors. This is a good example of how this study was limited by choice of records, for although great care was taken to find records exhibiting a predominance in harmonic structure none containing a clearly harmonic idea were found. There appear to be no adequate music pieces which are written in a purely harmonic idiom, and which are also recorded. Thus no harmonic factor was defined.

Another reason for this lack of a harmonic factor despite its prediction may be seen if one considers the aspect of familiarity of the listeners with the music. Introspectively, this writer would expect that all of the Vivaldi pieces played would score highly in a harmonic sense and would define this factor. However, it seems apparent that the untrained listeners heard even these, almost isorhythmic, pieces not as harmonic ideas, but as melodic ideas basically. This corresponds to the earlier comment about the reversal in variance between the melodic and the rhythmic factor. Perhaps, if the experiment were repeated a third time, and this idea of sophistication of listener response were valid, a harmonic factor would then be isolated.

experiment add other recordings which are introspectively similar enough so that some further definition of the factor may be achieved. This more refined choosing of the stimulus records would naturally lead to a better definition of the factors. Such a procedure of successive approximations, a technique used in all scientific endeavors, would eventually arrive at its ultimate goal, a clear, workable, and useful definition of musical factors.

## REFERENCES

1. APEL, W. *Harvard Dictionary of Music*. Cambridge: Harvard Univ. Press, 1950.
2. DRAKE, R. M. Factor analysis of music tests. *Psychol. Bull.*, 1939, **36**, 608-609.
3. CAMPBELL, I. G. Basal emotional patterns expressible in music. *Amer. J. Psychol.*, 1942, **55**, 1-17.
4. EYSENCK, H. J. The general factor in aesthetic judgments. *Brit. J. Psychol.*, 1940, **31**, 94-102.
5. GUILFORD, J. P., & HOLLEY, J. W. A factorial approach to the analysis of variance in esthetic judgments. *J. Exper. Psychol.*, 1949, **39**, 208-218.
6. GUNDLACH, R. J. An analysis of some musical factors determining the mood characteristics of music. *Psychol. Bull.*, 1934, **31**, 592-593.
7. HESTENES, M. R., & KARUSH, W. A method of gradients for the calculation of the characteristic roots and vectors of a real symmetric matrix. Washington: *National Bureau of Standards J. Res.*, July, 1951.
8. HEVNER, K. The affective value of certain elements of musical form. *Psychol. Bull.*, 1934, **31**, 678-679.
9. KARLIN, J. E. A factorial study of auditory function. *Psychometrika*, 1942, **7**, 251-279.
10. ———. Musical ability. *Psychometrika*, 1941, **6**, 61-65.
11. WING, H. D. A factorial study of musical tests. *Brit. J. Psychol.*, 1941, **31**, 341-355.
12. YOWELL, E. C. The determination of eigenvectors and eigenvalues of symmetric matrices. *Proceedings, Computation Seminar*. New York: I.B.M. Corp., 1951.

Department of Music  
University of California  
405 Hilgard Avenue  
Los Angeles 24, California



## MEMORY FOR COLOR\*

*Psychiatric Institute, Columbia University*

VIOLET HAMWI AND CARNEY LANDIS<sup>1</sup>

### A. THE PROBLEM

Everyday experience seems to tell us that some people have an excellent immediate memory for colors while others do not. Some persons after being shown a particular color seem able to bear the color in mind and to identify it or closely approximate it at a subsequent period while others cannot. A search of the literature provided us with only two experiments which seemed at all relevant to this point.

Rood (7) who utilized a color mixing wheel on which the mixtures of colors could be altered during rotation, and himself as the *O*, reported that after a glance at the color given by the mixture he could reset the disks immediately after the mixture had been changed, within an accuracy of one-half of one per cent. After an hour's lapse he could reset them with an average error of 2.2 per cent and after 24 hours, with an error of 4.5 per cent. Collins (1), using a spectroscope, exhibited four monochromatic spectral bands, yellow (588  $m\mu$ ), blue (460.9  $m\mu$ ), red (670  $m\mu$ ), and green (535  $m\mu$ ) to each of five *O*'s. Each *O* would view the color for 5 secs., look at a grey paper for a 15-sec. interval, and then attempt to reset the spectroscope, the setting of which had been changed during the interval, to the color originally seen. She found the yellow and blue were rather accurately reproduced while the red and green were not. Changing the green from 535  $m\mu$  to 500  $m\mu$  gave a color which was much easier to reproduce. Although pertinent to the problem in which we were interested, each of these experiments have the obvious limitations of a short delay between perception and recall, few colors, and few subjects.<sup>2</sup>

Preliminary investigation quickly demonstrated that the usual materials

\*Received in the Editorial Office on October 10, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was in part, made possible by a grant from the Carnegie Corporation of New York. The responsibility for statements and conclusions rests with the authors. We are most grateful for the assistance of Dr. Gertrude Rand and Dr. Joseph Zubin.

<sup>2</sup>A general description of a third experiment: "A color memory test," was reported in the *J. Opt. Soc. Amer.*, 1954, 44, 658-659, by R. W. Burnham and J. R. Clark. The complete detailed report has not yet been published, so no comparison of it with this study can be made.

and methods were not satisfactory for the purpose we had in mind, i.e., individual differences in memory for colors and the variables which affect them. Continuous variations of color by means of Polaroid color filters provided too many extraneous cues, while the usual available color materials such as papers or paints proved far too easy to disclose any differences between *O*s.

## B. MATERIAL

The experimental material necessary had to be such that it would have a large range of hues of differing saturation so that it would present some degree of difficulty for the *O*. Also it had to be made of a material which would not deteriorate or change with use.

A Color Harmony Manual (2) (1st Ed.) was secured which met the conditions which seemed necessary to us. This manual "consists of 12 handbooks containing 680 colored [plastic] chips representing an abridged Ostwald system [of color classification]. Each handbook presents two groups of 28 color chips, and each group forms one of the Ostwald triangles<sup>3</sup> [that is, by the addition of differing amounts of white or black to the basic hues]. The color chips within any group have an approximately constant dominant wave length. The other group shown in the same handbook has a dominant wave length approximately complementary to that of the first group" (5). The chips are made of transparent cellulose acetate to which pigmented film has been applied. Each chip has a glossy and dull surface. For this study the dull, unlettered side was used. The eight chips which made up the black to white series were not used in the experiment.

From the series of 672 chips we selected and used the following:

1 pe: Greenish Yellow; dominant wave length,	573.5	mμ (5)
21 nc: Green; dominant wave length,	499.2	mμ (5)
8 le: Purplish Red; dominant wave length,	494.0c <sup>4</sup>	mμ (5)
14 nc: Blue; dominant wave length,	475.7	mμ (5)
10 ni: Dark Purple; dominant wave length,	554.0c <sup>4</sup>	mμ (5)
6 ic: Pink; dominant wave length,	606.5	mμ (5)
17 pa: Greenish Blue; dominant wave length,	486.6	mμ (5)
3 gc: Yellowish Orange; dominant wave length,	582.9	mμ (5)
23 nc: Yellowish Green; dominant wave length,	550.8	mμ (5)
14 pa: Blue; dominant wave length,	475.4	mμ (5)

<sup>3</sup>Within each Ostwald triangle the arrangement is such that the chips going in a right and upward direction, have the black held constant while the amount of white is increased. When these same chips are examined going in a right and downward direction the amount of white is held constant and the black is increased. Each chip is tagged with a number and two letters. The number of the chip indicates its place in the hue circle, the first letter, *c* to *p*, indicates the relative amount of white (*c* denotes the largest amount); the second letter, *a* to *n*, indicates the relative amount of black.

<sup>4</sup>*c* denotes the complementary dominant wave length.

In the first experiment the first eight chips listed were used and were presented in the order named. In the second experiment seven chips were presented in the following order, 23 *nc*, 8 *le*, 21 *nc*, 10 *ni*, 17 *pa*, 6 *ic*, and 14 *pa*.

### C. OBSERVERS

Twelve normal healthy observers (7 female, 5 male), whose ages ranged from 22 to 50 years, took part in the experiment. They were all within the range of normal color vision as tested by the Farnsworth-Munsell 100 Hue test (4).

### D. METHOD

Each particular stimulus chip was presented to the observer on a large white matte surface cardboard placed on the top of a desk. This was illuminated by a Sylvania daylight lamp (A23) about 10 inches above the cardboard and shielded by an appropriate reflector. The room in which the experiments were conducted was illuminated by north light. No attempt was made to modify the outside daylight values, so that the illumination at the level of the *O*'s eyes varied from 10 to 20 foot-candles.

Two methods were used, an unmixed and a mixed. With the unmixed method the *O* was told: "Please look at this colored chip; try to isolate and label the color components and give a name to the color. Remember the color. In (15 minutes) (tomorrow morning) (at the end of the week-end) I will ask you to pick it out from the other colored chips." Each *O* looked at each chip 105 secs. This exposure time, determined during preliminary experiments, was of sufficient duration for visual inspection, without seeming unduly long. During the first three experimental days, the *O* returned at the end of 15 minutes and attempted to select the stimulus chip. On exposure Days 4, 5, and 6 the subject was required to return after 24 hours to make his selection. On Day 7 one-half of the *O*'s were asked to remember the chip for approximately 65 hours (Friday afternoon until Monday morning); the second half for 24 hours; while on Day 8 these two groups were reversed.

When the *O* returned, after the lapse of the appropriate time, he was read his description of the stimulus chip and asked to pick out the particular chip from the 12 books in which the chips were arranged in a systematic fashion. If his choice was not exactly right he was shown the original (correct) chip, so that he might compare the two; if his choice was correct he was so informed. This unmixed method was used with the first six *O*s in the first experiment.

After the first few applications of the unmixed method we were impressed

by the accuracy of the performance of the *O*s and so sought to make the procedure more difficult. In the hope of achieving this, we selected 168 chips (7 from each book) and placed them in a haphazard fashion on the white cardboard, but with a 6-inch diameter circle containing no chips, in the center of the scatter. The experimenter placed the experimental chip in the center of the vacant circle, giving the same instruction as with the unmixed method. When the *O* returned the 169 chips had been randomly moved about and the vacant circle was not apparent. He was asked to pick the appropriate chip from the mixed group of 169 chips. If he made an incorrect choice the correct choice was pointed out to him as before.

The entire experiment was then repeated a second time. In this second experiment the six *O*s who had used the unmixed method now used the mixed method and the other six, the unmixed. In the second experiment all *O*s, with both methods, were examined after a 15-min. interval. Three of the first experiment chips which were found to be relatively easy, were omitted (1 *pe*, 14 *nc*, 3 *gc*) and two were added (23 *nc* and 14 *pa*, a yellowish green and a blue, both having very similar black and white content as 21 *nc* and 17 *pa*), so shortening the series.

In summary the experimental design was as given in Table 1.

TABLE 1

Observers	No.	Experiment 1		Experiment 2	
		Method	Delay	Method	Delay
Group A	6	Mixed	15 min., 24 hr., 65 hr.	Unmixed	15 min.
Group B	6	Unmixed	15 min., 24 hr., 65 hr.	Mixed	15 min.

### E. SCORING METHOD

The scaling of differences between chips was not determined psychometrically, hence, we had no direct way of knowing how many *j.n.d.* steps existed between the chips in either of the three dimensions of hue, blackness, or whiteness. We adopted a simple but arbitrary scoring procedure. We assigned one error point for each step (in terms of the systematically arranged surround chips) away from the correct chip within its particular triangle for either blackness or whiteness. The errors in hue were counted in terms of the number of "book" step deviations. Faults arising from this scoring method will be considered below.

### F. RESULTS

The general accuracy of the memory for color under these conditions was surprising. For the two experiments the highest error score was 14, which



TABLE 2  
ERRORS MADE BY INDIVIDUALS AND ON DIFFERENT CHIPS UNDER VARIOUS CONDITIONS

O	Total			First experiment			Elapsed time			Un- mixed			Second experiment			White		
	Un- mixed			Black			White			Mixed			Hue			Black		
	Mixed	U	M	U	M	U	U	M	15'	24h	65h	U	M	U	M	U	M	U
MM	14	3		3			8		4	6	4	6	3			1		2
JF	14	6		4			4		6	7	1		5			3		2
FI	13	7		2			4		3	5	5		3			0		2
CI	12	4		4			4		5	4	3		1			1		4
CS	12	4		4			4		2	4	6		3			2		3
DL	7	2		2			3		5	2	0		3			1		2
NK					5			3	5	8	1					2		5
JC	10		2		4			4	1	6	3		6			5		2
CR	9		1		5			3	7	1	1		5			2		2
LS	8		4		2			2	5	2	1		4			4		4
CC	6		2		2			2	2	4	0		3			3		6
MB	2		1		0			1	2	0	0		2			1		3
Total	72	49	26	16	19	18	27	15	47	49	25	62	41	23	18	17	8	22
Chip																		
1 pe	3	4	1	0	0	1	2	3	7									4
21 nc	11	15	3	5	4	7	4	4	3	26						4		4
8 le	11	3	5	1	2	1	4	1	14							1		0
14 nc	5	2	2	1	1	0	2	1		7								
10 ni	10	4	6	1	1	1	2	3	1	14						1		2
6 ic	8	6	3	2	3	2	2	2	2	14						0		1
17 pa	22	11	6	4	8	4	8	3	3	13	20					4		7
3 gc	2	4	0	2	0	1	2	1		1	5					1		1
23 nc																4		4
14 pa																3		3
Total	72	49	26	16	19	18	27	15	47	49	25	62	41	23	18	17	8	22

was made by 3 *O*s. Their median error score for the individual chips was 2. Our most accurate *O* made but two errors in the entire first experiment, which errors were made on but one chip.

The results in terms of error scores on both the first and second experiments are summarized in Table 2.

### 1. Method

For both experiments it is apparent that the mixed method (which we had hoped would be more difficult) gave fewer errors than the unmixed. In the first experiment the total error score made when the unmixed method was used was 72 and that with the mixed was 49; in the second, the unmixed method gave 62 errors and the mixed 41. As may be noted the score made by 11 of the 12 *O*s with the unmixed method was equal to or greater than that made with mixed method.

### 2. Memory for Specific Colors

By referring to Table 2 it may be seen that the number of errors attributable to each chip is different. The two chips giving rise to the most errors were 17 *pa* (greenish blue) and 21 *nc* (green). In the first experiment they accounted for 46.2 per cent of the errors made with the unmixed method and 53.0 per cent of those with the mixed. In the second experiment these two chips were responsible for 45.9 per cent of the unmixed and 43.9 per cent of the mixed method errors. A tabulation showed that most individuals had their higher error scores on Chips 17 *pa* and 21 *nc*.

Although the color vision of all *O*s was in the normal range as indicated by the Farnsworth-Munsell 100-Hue Test, some subjects gave a more accurate performance than others. In order to determine whether higher error scores attributable to 21 *nc* and 17 *pa* might be related to some color weakness, a comparison was made of the error scores on these chips for the four of the *O*s who had the best performance in the blue-green area on the 100-Hue Test and the four *O*s having the poorest performance. This comparison showed that the total number of errors made on Chips 21 *nc* and 17 *pa* was exactly the same, 33, for the four best and four poorest, showing that the difficulty in identifying these two chips could not be attributed to weakness of color vision.

The chips varied among themselves in the qualities of hue, blackness, and whiteness. Hence, we tabulated the total error points that were made with respect to each quality, together with the direction of the errors and the way in which the individual chips contributed to these errors. These results are

given in Table 3 where the ordering of the chips is presented in terms of the total error score. (With respect to hue, + means an error in the violet direction, — an error in the red direction). From Tables 2 and 3 the following statements may be made.

TABLE 3

MEMORY ERRORS MADE ON EACH CHIP USED, GROUPED WITH RESPECT TO HUE, BLACKNESS, AND WHITENESS, TOGETHER WITH THE DIRECTION OF THE ERRORS  
(Asterisk signifies that chip was used in only one of the experiments)

Chip	Hue		Blackness		Whiteness		Total
	+	—	+	—	+	—	
17 pa	21	0	19	0	19	0	59
21 nc	5	12	14	1	8	7	47
8 le	4	7	4	1	3	6	25
10 ni	14	1	3	1	4	2	25
6 ic	7	2	3	3	2	3	20
14 pa*	.0	3	5	0	7	0	15
23 nc*	0	1	1	4	1	6	13
1 pe*	1	0	1	0	4	1	7
14 nc*	2	1	1	0	3	0	7
3 gc*	1	1	1	0	0	3	6
Total	55	28	52	10	51	28	224
% of Total	24.6	12.5	23.2	4.5	22.8	12.5	100.1

There was no tendency for all the chips to give a preponderance of errors in one quality or another or in one direction or another. Chip 17 *pa* was a "fully saturated" hue and could give errors only in the + direction with respect to blackness or whiteness.

In general, color memory tended to gain in blackness. This was true even when the errors attributed to Chips 17 *pa* and 21 *nc* were omitted. It was shown during either method and with any of the three delay periods. The apparent average shift in color memory toward the violet end of the spectrum depended largely on the errors attributable to Chips 17 *pa* and 10 *ni*; the remaining chips showed no such constant tendency.

That individual differences in accuracy of color memory exist, seems very obvious from Table 2. However, if the differences were to be attributed to any single factor such as memory, per se, then approximately the same rank order among *O*s should prevail as any single variable was changed in these experiments. This would be indicated by a high rank difference correlational matrix. We determined these correlations with respect to the two experiments, two methods, the 15-min. and 24-hour time lapse, hue-blackness-whiteness and specific chips. All these correlations were low and statistically not significant. The conclusion must be drawn that the apparent individ-

ual differences were the result of the various combinations of conditions used and, so far as our evidence goes, not particularly dependent on the *O* himself.

When a comparison of the errors was made on those individual chips which were used in both experiments, it was found that in 8 out of 10 instances the first error score exceeded the second error score. This difference is significantly greater than chance and indicates that some practice effect was carried over from one experiment to the next. Within either single experiment, the gross differences in memory accuracy for the individual chips overshadowed any evidence of practice effect which may have occurred.

In the first experiment we varied the time lapse between perception and recall. However, the differences between the difficulty of the individual chips makes any evaluation of the influence of the time lapse on memory accuracy, difficult. One of the three chips used for the 15-min. interval was 21 *nc* which chip gave rise to a disproportionate number of errors. Half of the *O*s were asked to remember 17 *pa* for the 65-hour interval while the other half of the *O*s were asked to remember 3 *gc*. It happened that 17 *pa* gave rise to 80 per cent of all errors during these conditions, while 3 *gc* gave rise to only 20 per cent. Our sample was too small to provide a basis for any decisive conclusion but certainly the indication is clear that time lapse was not as important a factor as was the difference in difficulty attributable to the individual chips.

#### G. OBSERVERS' COMMENTS

At the end of the second experiment each *O* was asked which method he preferred and why. Seven *O*s reported that the mixed method seemed easier; 6 of these gave as their reason that they could compare the stimulus chip with the surrounding chips or use the surrounding chips as reference points, while the seventh *O* stated that finer discriminations had to be made on the unmixed method. All of these 7 *O*s gave a more accurate performance with the mixed method. Four *O*s reported that they found the unmixed method easier. Observer *NK* felt that in the unmixed method there was "more remembrance of color—mixed was more dependent on description and not as sure of the color—would forget the color." *NK* was the only *O* who did better with the unmixed method than with the mixed. Observer *MM* stated that the "mixed was more work because I had to hunt for possibilities and I felt I might not have included the stimulus chip in the possibilities." *JF* considered the unmixed easier because, "It's easier to fix the color without distraction and easier to pick it out because of the allied colors in the book."



*DL* gave no reason for preferring the unmixed method. *CC* reported no preference as to method stating that "neither is easy, I never had a sure feeling about any of the chips. I always felt as if I were guessing."

## H. DISCUSSION

Several points deserve further comment; the scoring of the errors, the gross differences in ease of memory for the different colors, the accuracy of immediate memory for color, and the effect of the two methods used, on the results obtained.

We exhibited a small plastic chip having distinctive, definable qualities of hue, blackness, and whiteness under controlled conditions. After a time lapse our *O*s were asked to select this particular chip from either an ordered, unmixed array of 672 chips or a scattered mixed group of 169 chips. We obtained an error score in terms of each quality by counting the number of chips (or books) the selected chip was from the correct chip. If we had had information as to the probable number of discriminable differences in each of the three qualities which could be seen by the "average eye" under standard conditions, between the exposed chip and the chip picked after a time lapse, our scoring would have been more rational.

The methods used and the difficulties inherent in constructing a color scale have been reviewed recently by Judd (6). Seemingly scaling procedures can provide agreed upon steps of equal discriminable difficulty for the color solid under certain specified conditions and assumptions. Methods giving smaller step units than we employed have been utilized but circumstances did not permit us to use such methods. We rather doubt that the application of some tristimulus analysis anchorage and scoring would have changed the direction of the results we have reported although it is probable that a better statistical evaluation might have resulted. In order to clarify several points mentioned in the results an improved method and scoring system should be used in further experiments.

In the selection of the stimulus chips an attempt was made to choose hues which were not close together in the spectrum. As noted there were gross differences in the recognition of these chips, with Chips 21 *nc* and 17 *pa*, a green and a greenish-blue giving rise to almost half of all the errors. Furthermore the error pattern associated with each chip seemed distinctive for that chip. Collins (1) found the spectral green of 535  $m\mu$  difficult for her subjects to reproduce while 500  $m\mu$  was easy. She concluded that particular spectral wave bands "possess some qualities inherent in themselves which make them difficult to recognize again and almost impossible to reproduce." We con-

cur. However, it does not seem that hue alone is responsible for the greater or lesser difficulty of immediate memory for color. A hint of this is given by the results we obtained with the two blue chips, 14 *nc* and 14 *pa*, which have very similar dominant wave lengths, 475.7 and 475.4. The first, 14 *nc*, yielded very few errors while 14 *pa*, as used in the second experiment, proved to be of moderate difficulty. Hence, in addition to Collins' finding (and ours) that a change in hue is sufficient to change the difficulty of the recognition of a color, we have also shown in this study that the amount of blackness and/or whiteness are involved. Why there should be such great differences in the immediate memory value of specific hues and saturation of hues deserves further investigation.

In spite of these marked differences in accuracy of memory for certain specific colors it is still remarkable how few the errors in recognition were and how often the memory was exact. Each of our 12 *O*s attempted 15 identifications after various time lapses. Out of the 180 attempts 63 were exactly correct.

The remaining 117 incorrect choices make up a total error score of 224. Of these there were 83 errors for hue, 62 for blackness, and 79 for whiteness. The lesser score for blackness is probably not significant but the fact that it occurred in the first experiment with the unmixed condition and in the second experiment under both conditions is interesting. Another peculiarity in the judgment of blackness, is that when an error was made it was more often in the direction of greater blackness rather than of less blackness.

When we compared individuals with respect to immediate color memory we found differences but these differences seemed for the most part to grow out of the variations we introduced in these experiments. Collins (1) had one subject who "developed a pure memory of the yellow" (588 *mμ*), being able to reproduce it after a two-month lapse with an error of but 0.1 *mμ*. In our first experiment *MB* made only two errors, both of these on the identification of 21 *nc* (picking 20 *pc* instead) after a 15-min. lapse and no errors on any of the other chips some of which involved time lapses of 24 hours or 65 hours. At the opposite extreme were three *O*s each of whom made 14 errors on the first experiment. Of the 14 errors made by each, 5, 5, and 6 of the errors were made on Chips 17 *pa* and 21 *nc*. In the second experiment the high error score was made by one *O*, who was not among the three *O*s of the first experiment. He made 13 errors, including four attributable to 17 *pa*. That each subject was read his own description of the stimulus chip before he made his choice may be an important factor giving rise to this relatively high degree of accuracy. There were instances,

however, when the description was not particularly accurate or where certain components were overemphasized which may have tended to mislead rather than aid the *O*. Further research on this point is indicated.

Finally, the difference in the error scores given by the two methods may be simply the result of the number of chips from which the *O*s had to select, namely 672 or 169. This seems unlikely since the number of the mixed chips was but a quarter of the unmixed number while the mixed error score was only a third less than that produced by the unmixed method. A more probable explanation is that given by the *O*s, namely, that the presence of the other chips, at the time of memorization enabled them to have a better knowledge of the chips which might more easily give rise to confusion.

### I. SUMMARY

The immediate memory for color has been studied with 12 *O*s in two experiments. Ten plastic color chips from the *Color Harmony Manual* (1st Ed.) served as stimuli. In the unmixed method, the *O* was required to identify the stimulus from the entire 672 chips systematically arranged; in the mixed method he was required to select it from among 169 chips arranged in a haphazard fashion. The variables studied were hue, blackness, whiteness, uniqueness of particular chips, individual differences among *O*s, method of presentation for recognition, time lapse, and practice.

The relative exactness of immediate color memory and the ease or difficulty of exact recognition of a particular stimulus, were the outstanding findings. These two factors were so marked that the effects of the other variables were obscured and could only be considered as possible tendencies. These tendencies seemed to indicate that the length of the time lapse, whether 15 min., 24 hours, or 65 hours made little difference; individual differences between *O*s were obscured by the combinations of the other variables, hue, blackness or whiteness contributed in a chance fashion to error scores and only little practice effect was shown. The unmixed method did give rise to more errors than the mixed method.

### REFERENCES

1. COLLINS, M. Some observations on immediate colour memory. *Brit. J. Psychol.*, 1931-32, **22**, 344-352.
2. COLOR HARMONY MANUAL. Chicago: Container Corp. of America, 1942.
3. FARNSWORTH, D. The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. *J. Opt. Soc. Amer.*, 1943, **33**, 568-578.
4. ———. Farnsworth-Munsell 100-Hue Test for the Examination of Anomalous Color Vision. Baltimore, Md.: Munsell, 1943.

5. GRANVILLE, W. C., & JACOBSON, E. Colourmetric specifications of the Color Harmony Manual from spectrophotometric measurements. *J. Opt. Soc. Amer.*, 1944, **34**, 382-395.
6. JUDD, D. E. Color in Business, Science, and Industry. New York: Wiley, 1952. (Pp. 170-200.)
7. ROOD, O. N. Our memory for colour and luminosity. *Nature*, 1879, **21**, 144.

*Psychiatric Institute*  
*722 West 168th Street*  
*New York City 32*



## "FEARS AND WORRIES" CONCERNING PHYSICAL CHANGES: A PRELIMINARY SURVEY OF 32 FEMALES\*

*College of Education, University of Oklahoma*

HENRY ANGELINO AND EDMUND V. MECH<sup>1</sup>

### A. THE PROBLEM

One of the major tasks of the growing individual is to come to terms with his own developing body, which means learning to adjust to the often very dramatic changes in his physical appearance.

Equally important is the not too clearly defined rôle of the classroom teacher in such situations. Yet modern educational practice supports the view that every aspect of the child's growth must be given attention. In this framework, a teacher is required to pay attention to the social and emotional needs of the pupil as well as those circumscribed by the intellectual dimension. Not any less important is parental understanding of growth phenomena. Simply then, the present study was oriented toward the gathering of "fears and worries" relative to problems concerning physical appearance and development.

It was felt that these data would be especially valuable to teachers and parents who, to quote Horrocks (2, p. 261), "may do their adolescent charges considerable service if they will prepare them for physical change. . . ."

### B. PROCEDURE

Thirty-two females, mean age 20.3 years, enrolled in a "Human Growth and Development" course at the University of Oklahoma were requested to—"write a brief anonymous account of some aspect of physical development which disturbed you either during childhood or adolescence." Subjects were allowed complete freedom in framing their reactions to the foregoing request. College students were used because it was felt they were sufficiently close to these changes to be accurate in their reporting as well as being somewhat more sophisticated than the usual high school student upon whom most of the published studies were made.

In a recent paper Strang (6) indicates that such "introspective" devices may indeed help fill some gaps in our knowledge of adolescent behavior. For

\*Received in the Editorial Office on October 10, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>We are indebted to the Faculty Research Committee of the University for financial support.

studies utilizing young children with a slightly different orientation the reader is referred to (1, 3, 4, 5, 7).

### C. RESULTS AND DISCUSSION

The results of the preliminary survey are shown in Table 1. The variety of responses obtained were aptly classified into five "fears and worries" categories, as follow: Height, Weight, Teeth Defects, Speech Defects, and Fear of Puberty and Menstruation.

TABLE 1  
CLASSIFICATION AND FREQUENCY OF THE "FEARS AND WORRIES" OF 32 FEMALES  
CONCERNING PHYSICAL CHANGES

Fear or worry	Frequency of occurrence
A. Height:	
Tallness	11
Shortness	1
B. Weight:	
Underweight	3
Overweight	7
C. Teeth defects	3
D. Speech defects	2
E. Menstruation, puberty	5
Total	32

Concerns with height and weight dominated the worries of physical changes cited by the 32 females. Of the 32 females sampled, approximately 38 per cent expressed height as causing the major physical problem to them. Weight followed a close second in terms of causing concern to the physical "self"; being reported by 31 per cent of our group. Concern over puberty and menstruation, teeth and speech defects, constituted the remaining fears and worries respectively.

Numerical tabulations, however, often tend to mask the important underlying contributions that are embodied in the data. In the present investigation many of the "protocols" proved extremely illuminating and for this reason merit exposition. Verbatim excerpts from several protocols follow:

*Case 1:* "I still grit my teeth when I think of the agony I thought I was going through because of my height. I reached my maximum height when I was a freshman in high school, and I was a half-inch taller than my brother, six years older. Friends would look at me and say, 'My, but you're taller than your mother.' I would look down at all five feet of her and say, 'Yes, I guess I am, ha ha.' I wondered if I had been born five feet, eight inches tall."

*Case 2:* "I was taller than most girls my age and taller than nearly all the boys. I felt awkward and out of place. I resented always having to be on the back row when a picture was taken. As I grew older I avoided going to parties and dances because I felt the boys preferred short girls with more rounded figures."

*Case 3:* "Irregular teeth represent the biggest problem I had in physical development during childhood. As soon as it was evident that my permanent teeth left much to be desired as far as perfect alignment was concerned, it seems to me that all relatives and acquaintances began to 'harp' at my parents (*in front of me*) to have something done about them. I became very self-conscious about them. I even dreaded meeting new people."

*Case 4:* "When I was in junior high school I was very conscious of being skinny. My classmates didn't tease me about it very often but I felt they were just being nice. My best girl friend was what the boys called 'stacked' and this only caused me to feel worse."

*Case 5:* "Suddenly, at the age of about eleven, everything I ate seemed to turn into fat. I became more than just 'pleasingly plump.' The only possible thing you could say about me was that I was just 'plain fat.'"

"Naturally, having always been more of the average in size, everyone noticed my increase in weight. Most of them were kind enough to make some comment on this fact of which I was already all too well aware."

"Their comments did little to help me overcome my self-consciousness. In fact, they made me feel even more self-conscious of my gain in weight."

*Case 6:* "One of the most pressing problems I was most aware of while growing up was adolescent overweight. I believe, being a girl, I was even more self-conscious about it than a boy would have been. During junior high school it was worse than any other time. I was not extremely fat, but was overweight enough to be aware of it. I believe this led to a feeling of social insecurity in my mind, all of which has not left me yet."

*Case 7:* "There was no phase of my development that worried me like the beginning of the menstrual period. I began this stage of growth at the age of ten, and it caused me a great deal of embarrassment for two reasons; first it was never explained fully to me and I did not understand its purpose nor did I realize what the process involved. I was also confused because none of the other girls my age had begun this period. I was so embarrassed about it that I would not ask my Mother to explain it to me, and due to this ignorance I believed for a long while that this function must be limited to my family alone since I knew that Mother went through the same cycle."

*Case 8:* "The aspect of physical growth that disturbed and embarrassed me in my adolescence was connected with the first evidence that

I was becoming a woman. These occurred when I was still quite young and was visiting with relatives. I woke up one morning with dark brown stains on my pajamas. Of course, I had no idea what was up, nor did I feel ill in any way. Soon after that Mother explained the 'facts of life' to me, but it was at least a year later before anything happened again. I was in the sixth grade then, and I had to 'sit out' during gym. This embarrassed me more than anything else, as I was sure that everyone knew just what was going on and I felt terribly self-conscious about it."

These preceding excerpts adequately point up several factors that bear import for education.

In not a single case did a respondent indicate a concern relating physical changes to schoolwork in any academic or intellectual sense. Rather, the emphasis was upon emotional or social concomitants.

Our data suggest that pupils exhibit a high proportion of fears and worries that are *not* associated in any direct manner with academic work in the school-room.

Certainly, in order to engender the mental health movement in the schools, the modern teacher must be sensitive to the multiplicity of needs that exist among school children on dimensions other than the intellectual. In effect, the teacher must become a human engineer by virtue of adjusting to the pupil.

#### REFERENCES

1. FRAZIER, A., & LISONBEE, L. K. Adolescent concerns with physique. *Sch. Rev.*, 1950, **58**, 397-405.
2. HORROCKS, J. E. *The Psychology of Adolescence*. Boston: Houghton Mifflin, 1951. Pp. 614.
3. JERSILD, A. T., & HOLMES, F. B. Children's fears. *Child Devel. Monog.*, 1935, No. 20.
4. JERSILD, A. T., MARKEY, F. V., & JERSILD, C. L. Children's fears, dreams, wishes, daydreams, likes, dislikes, pleasant and unpleasant memories. *Child Devel. Monog.*, 1933, No. 12.
5. PRATT, K. C. A study of fears of rural children. *J. Genet. Psychol.*, 1945, **67**, 179-194.
6. STRANG, R. Adolescents' views on aspects of their development. (Abstract of paper delivered at the 1954 meetings of the American Psychological Association.) *Amer. Psychol.*, 1954, **9**, 478-479.
7. STOLTZ, H. R., & STOLTZ, L. M. Adolescent problems related to somatic variations. In *Adolescence*, 43rd Yearbook, N.S.S.E. Chicago: Univ. Chicago Press, 1944.

College of Education  
University of Oklahoma  
Norman, Oklahoma



# CONTRIBUTIONS TO THE STANDARDIZATION OF THE TOMKINS-HORN PICTURE ARRANGEMENT TEST: PLATE NORMS\*<sup>1</sup>

*Department of Psychology, Princeton University*

SILVAN S. TOMKINS AND JOHN B. MINER

## A. PROBLEM

Notwithstanding the significant contribution of projective tests to our knowledge of personality, limitations arising from inadequate sampling procedures have been the rule rather than the exception. This has been a consequence in part of the time required to administer and score projective tests. The Tomkins-Horn Picture Arrangement Test<sup>2</sup> was devised to maximize ease of administration and scoring at the least cost in richness of projective material. The rationale of the test has been described elsewhere (4). In this investigation we have attempted to integrate the sampling procedures of social psychology and the projective methods of clinical psychology.

A representative normal sample of the United States population will be systematically compared with a pathological sample matched for those background factors which show a consistent effect on *PAT* performance—age, education, and intelligence.

This is the first of a series of papers devoted to the presentation of this normative data. In this paper we will present the distribution of arrangements for the 25 individual plates of the test based on a representative sample of 1,500 cases drawn from the non-institutionalized population (over nine years of age) of the United States. A second publication will present frequency distributions for a series of keys which have been developed by using groups of responses to a number of plates. Later publications will present similar frequency distributions by plates and sequences of plates for a representative sample of neurotic and psychotic records.

## B. PROCEDURE

This sample was collected by the interviewers of Public Opinion Surveys, Inc., of Princeton, New Jersey, using a sample design similar to that em-

\*Received in the Editorial Office on October 12, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was supported by the Medical Research and Development Board, Office of the Surgeon General, Department of the Army, under Contract No. DA-49-007-MD-476: OI No. 7-54.

<sup>2</sup>Hereinafter referred to as the *PAT*.

ployed in the national surveys conducted by that organization. The design provides stratification by seven regions (groups of states), and within each region stratification by geographical distribution of the population, three urban-rural strata, United States Census economic areas, and size of locality. A systematic sample of localities was drawn within each stratum from a random start with probability of selection proportional to size. Within large urban communities sampling units (small clusters of blocks) were drawn at random with probability proportional to size. In smaller communities and rural regions, sampling areas were drawn with equal probability. Interviewers were assigned selected areas or clusters of blocks and were required to work within the boundaries of such areas, where they chose respondents on the basis of quota assignments by sex and age. This phase of the project was carried out during the months of August and September, 1953.

When the total 1,500 interviews had been returned, the answers to cross-section questions were checked against the latest available Census data and the sample compared with population figures. The distribution of respondents was found to be inadequate in certain categories, and accordingly a supplementary sample design was set up employing a quota system for the selection of respondents on the basis of sex, education, and more specific age categories instead of the original area sampling with quota control. The prior stratification in terms of geographical, urban-rural, and economic areas as well as size of locality was, however, continued. Quotas were assigned in such a way as to fill out deficiencies in the original sampling. The discarding of the area sampling procedure was necessitated by the specificity required in selecting respondents falling in certain very narrow educational and age groupings. Interviewing on this phase was conducted during December, 1953, and January, 1954. In all 1,896 completed protocols were obtained, from which a representative sample of 1,500 cases was selected. The survey included 329 sampling areas in 228 localities (cities, towns, and counties). In each sampling area from five to seven interviews were assigned.

The standard interview consisted of the *PAT*, a 20-word multiple-choice vocabulary test, and a group of cross-section questions. The time required to complete these varied from 45 minutes to 2½ hours per respondent. In a few cases the interviewer was unable to get responses to all 25 *PAT* plates on the first call and had to return at a later date to complete the interview. Only complete protocols were accepted for the final analysis.

The interviewer introduced himself as follows:

I'm an interviewer for the Gallup Poll. We're making a nation-

wide survey for a group of leading psychologists on a picture-arrangement experiment. Is there a table inside where we can go over this together?

After being seated the interviewer opened the test booklet to the sample page and gave the following instructions.

This is an experiment to see how well you can arrange pictures so as to tell a story that makes sense. Each of the three pictures has a mark, or symbol, at the bottom. One is marked with a triangle, like that one (pointing to the triangle at the bottom of the picture on the sample page with that symbol), another is marked with a circle, like that one (pointing to the circle at the bottom of the picture on the sample page with that symbol), and the other is marked with a rectangle, like that one (pointing to the rectangle at the bottom of the picture on the sample page with that symbol). For each page, give me the symbol of the drawing which should come first, then the symbol of the drawing which should come second, and then the symbol of the drawing which would come last. Then, tell me the story for that page, using one sentence for each drawing. You don't have to write anything. . . . I'll take care of that.

In a few cases (0.9 per cent) the respondents insisted on writing down their own responses. The interviewers were instructed to take down the responses verbatim and not to assist the subjects in any way in interpreting the pictures. The only exception was in the case of illiterate subjects where it was necessary to read the very limited amount of written material included in the pictures. Occasionally a respondent would use the same symbol arrangement for a number of plates or adopt a position preference, such as always going from top to lower left to lower right. (The pictures are shown at angles of 120 degrees apart so that no particular sequence of drawings is favored.) In such cases it was explained that the arrangement which makes the best sense may vary from plate to plate, and the subject was requested to start again. Occasionally these records slipped by the interviewers, but every case of symbol and position preference was removed from the final sample.

The vocabulary test, an abbreviation of the vocabulary section of the *I.E.R.* Intelligence Scale, *CAVD* consisting of 20 words with five multiple-choice alternatives for each and developed by Thorndike and Gallup (2, 3), was introduced next as follows:

We're trying to find out how familiar these words are to people. Will you please look at this card and give me the number of the word that seems closest to each word in capital letters.

Again responses were recorded on an answer sheet. In the case of illiterate subjects the test was given orally. If the respondent obviously had difficulty in using the English language due to a lack of exposure to it and spoke some other language fluently, no intelligence score was computed. All subjects were urged to guess if they did not know a word. If a test was not completed, a correction for not guessing was added to the total score according to the number of words skipped. Utilizing a chance expectancy, we added nothing if two or less words were left out, one for three through seven, two for eight through 12, three for 13 through 17, and four for 18 or more. In 92.7 per cent of the protocols no correction of any kind was necessary.

The interview was terminated with a series of cross-section questions consisting of the last grade completed in school, occupation, occupation of the head of the family, marital status, religious preference, age, race, sex, rural-urban residence, and subjective class identification. This latter was determined on the basis of the question wording developed by Centers (1).

Data from the most recent Current Population Reports issued by the Bureau of the Census were used in setting up expected frequencies for analyses of each of seven background variables, as indicated in Table 1. The 1950 census formed the basis for rural-urban frequencies, since there is no more recent data available on the population of cities and towns throughout the country. Using *IBM* sorting procedures, a sample of 1,500 was drawn from the total of 1,896 cases in such a manner as to approximate most closely the expected frequencies on these eight variables. No comparable data on the total population is available on religion, occupation of family head, subjective class identification, and intelligence.

The categories employed in classifying the background characteristics of the respondents were determined according to the following specifications:

*Education.* Classification was based on the last grade completed. Respondents who had been trained in business, trade, or beauty schools, as well as nurses whose education was received outside of a regular academic college were considered as high school graduates. Junior college education was taken as equivalent to an incomplete college course.

*Religion.* The question on religious preference was the only one asked. No attempt was made to get at church membership, attendance, etc.

*Occupation of Family Head.* The chief wage earner of the family was in general considered as the family head. In some cases this was the respondent and in other cases it was not. Classification was determined through the use of the Alphabetical Index of Occupations and Industries prepared by the Bureau of the Census.



*Occupation of Respondent.* Employed categories were determined as with the heads of family, except that certain groups were combined so as to achieve large enough frequencies to give stable normative data. "Unemployed" is understood to include only those who are actively looking for a job, who are in the labor force.

*Class.* The respondent's own opinion was accepted. He was offered a choice of the four categories and was to choose one.

*Rural-Urban.* The "Farm" classification was restricted to those respondents who actually lived on farm property. All others who lived in the country and those living in towns of under 2,500 population as of 1950 were placed in the "Under 2,500" group.

*Geographical.* Grouping by state was carried out as follows: *New England* (Me., N. H., Vt., Mass., Conn., R. I.); *Middle Atlantic* (N. Y., N. J., Pa., Del., Md., W. Va.); *East Central* (Ohio, Mich., Ill., Ind.); *West Central* (Wisc., Minn., Iowa, Mo., N. Dak., S. Dak., Nebr., Kans.); *Deep Border South* (Va., N. C., Ky., Tenn., Okla., Texas, Fla., D. C.); *South* (S. C., Ga., Ala., Miss., La., Ark.); *Rocky Mountain* (Mont., Ariz., Colo., Idaho, Wyo., Utah, Nev., N. M.); and *Pacific Coast* (Calif., Ore., Wash.).

*Intelligence.* The five intelligence groups were selected according to the schema developed by Wechsler (5) to approximate the percentages of the population indicated in Table 1 under "Population Percentages." In order to correct for age, the total sample was divided into two-year intervals (10 and 11, 12 and 13, etc.) and distributions of vocabulary scores were set up for each group. Then each distribution was separated into five groups approximating as nearly as possible an 8.9, 16.1, 50.0, 16.1, 8.9 percentage break. When these cutting points had been established, a smoothed curve was fitted by eye in an attempt to eliminate random sampling fluctuations. Thus, at age 10 a score of 13 or more correct is considered superior, while at 45 the respondent must get at least 18 right to reach the same level of performance.

Statistical analysis of the deviation of our sample frequencies from those expected on the basis of data available for the total population indicates that in no case does a sample frequency differ significantly from the population value. In Table 1 we have not indicated the number of respondents for whom data are not available on a particular background variable, and therefore, sample percentages do not always total 100 per cent.

TABLE 1  
SAMPLE COMPARED WITH POPULATION DATA

Background variables	Sample N	Sample per- centages	Population per- centages	Background variables	Sample N	Sample per- centages	Population per- centages
<i>Education</i>				<i>Occupation of respondent</i>			
0-6	309	20.6)	45.0	Professional,			
7-8	367	24.5)		Managers &			
9-11	296	19.7	19.9	Farmers	185	12.3	12.7
12	333	22.2	22.3	Clerical &			
13-15	107	7.1	7.1	Sales	135	9.0	10.0
16+	88	5.8	5.7	Skilled	98	6.5	6.8
<i>Sex</i>				Operatives	156	10.4	10.3
Male	724	48.3	47.9	Service &			
Female	775	51.7	52.1	Laborers	164	10.9	9.9
<i>Marital status</i>				Students	251	16.7	16.7
Married	941	62.7	63.1	Unemployed	20	1.3	1.5
Single	410	27.3	27.3	Housewives	433	28.9)	
Widowed	121	8.1	7.8	Retired	57	3.8)	32.1
Divorced	28	1.9	1.8	<i>Class</i>			
<i>Religion</i>				Upper	46	3.1	
Protestant	1,120	74.7		Middle	586	39.1	No
Catholic	294	19.6	No	Working	785	52.3	Data
Jewish	49	3.3	Data	Lower	61	4.1	
Other	37	2.5		<i>Rural-urban</i>			
<i>Age</i>				Farm	229	15.3	15.3
10-13	122	8.1	8.2	Under 2,500	382	25.5	25.9
14-17	113	7.5	7.2	2,500-9,999	153	10.2	9.7
18-24	164	10.9	10.4	10,000-49,999	207	13.8	13.7
25-34	281	18.7	19.0	50,000-499,999	265	17.7	17.7
35-44	264	17.6	18.0	500,000+	264	17.6	
45-54	225	15.0	14.8	<i>Geographical</i>			
55-64	174	11.6	11.6	New England	93	6.2	6.1
65+	152	10.1	10.8	Middle			
<i>Race</i>				Atlantic	342	22.8	22.7
White	1,352	90.1	89.9	East Central	271	18.1	18.1
Negro	148	9.9	10.1	West Central	178	11.9	11.3
<i>Occupation of Family head</i>				Border South	266	17.7	18.1
Profession	123	8.2		Deep South	150	10.0	9.9
Farmers	180	12.0		Rocky Mountain	53	3.5	3.5
Managers	116	7.7		Pacific Coast	147	9.8	10.3
Clerical	111	7.4	No	<i>Intelligence</i>			
Sales	103	6.9	Data	Subnormal	131	8.7	8.9
Skilled	245	16.3		Low Average	239	15.9	16.1
Operatives	309	20.6		Average	751	50.1	50.0
Service	131	8.7		High Average	237	15.8	16.1
Laborers	148	9.9		Superior	135	9.0	8.9

TABLE 2  
RESPONSE FREQUENCIES FOR THE *P-IT* EXPRESSED AS PERCENTAGES

Back-ground Vari- ables	Arrangements					Back-ground Vari- ables	Arrangements						
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV
Plate 1													
Total sample	27.3	27.7	4.1	11.8	23.0	6.1	Education						
Intelligence							0- 6	22.7	28.2	4.5	15.6	22.0	7.1
Subnor- mal							7- 8	23.9	26.7	3.8	13.9	22.6	9.0
Low Av.	22.9	23.7	4.6	18.3	19.8	10.7	9-11	33.1	24.7	3.4	11.2	22.0	5.7
Average	25.1	29.3	5.0	12.6	20.5	7.5	12	28.5	29.4	5.4	7.8	25.8	3.0
High Av.	27.3	27.3	4.0	11.4	24.6	5.3	13-15	25.2	33.7	2.8	9.4	23.4	5.6
Superior	30.0	30.8	3.8	8.9	22.4	4.2	16+	35.2	27.3	3.4	10.2	20.4	3.4
	30.4	26.7	3.0	10.4	23.0	6.7							
Age													
10-13	22.1	16.3	4.1	19.7	29.5	8.2	35-44	30.3	29.6	3.4	11.0	20.1	5.7
14-17	38.0	15.0	3.5	11.5	21.2	10.6	45-54	23.6	28.9	5.3	9.8	24.4	8.0
18-24	35.4	28.7	2.4	6.7	22.6	4.3	55-64	23.0	33.9	1.7	14.4	19.6	7.5
25-34	28.8	28.1	5.0	8.2	26.0	3.9	65+	17.1	32.9	6.6	19.1	21.1	3.3
Plate 2													
Total sample	10.5	4.3	10.9	2.5	11.2	60.8	Education						
Intelligence							0- 6	8.1	7.1	15.9	5.2	14.6	49.2
Subnor- mal							7- 8	8.2	5.4	13.1	3.3	9.5	60.4
Low Av.	6.9	5.3	25.2	6.1	11.4	45.0	9-11	10.5	2.7	11.5	2.0	11.2	62.2
Average	9.2	6.7	13.4	3.3	10.0	57.3	12	12.6	3.0	6.9	0.9	9.6	66.9
High Av.	10.4	4.0	10.0	2.3	12.1	61.2	13-15	13.1	1.9	6.5	0.0	14.0	64.5
Superior	10.6	3.0	6.3	0.8	9.7	69.6	16+	17.0	2.3	2.3	0.0	9.1	69.3
	14.8	3.0	5.2	0.7	10.4	65.9							
Age													
10-13	12.3	5.7	18.9	2.5	4.1	56.6	35-44	11.0	4.9	8.7	2.3	11.0	62.2
14-17	13.2	4.4	8.9	0.0	8.9	64.6	45-54	7.6	6.2	10.7	1.3	11.6	62.7
18-24	13.2	4.4	8.9	0.0	8.9	64.6	55-64	8.1	8.1	13.8	2.3	16.1	51.8
25-34	12.2	1.3	10.4	1.2	9.8	65.2	65+	10.5	0.7	13.8	7.9	18.4	48.7
	10.3	2.8	7.5	2.1	8.9	68.4							
Plate 3													
Total sample	29.9	2.5	21.3	20.1	7.4	18.8	Education						
Intelligence							0- 6	27.5	3.9	18.8	20.4	9.4	20.1
Subnor- mal							7- 8	25.9	3.3	22.0	22.3	7.6	18.8
Low Av.	29.8	3.8	20.6	20.6	7.6	17.5	9-11	33.8	1.4	21.6	19.6	6.4	17.2
Average	25.1	4.2	19.7	25.1	8.8	17.2	12	31.2	2.4	22.5	17.4	7.2	19.2
High Av.	31.3	2.3	21.9	18.8	7.0	18.6	13-15	23.4	1.9	26.2	22.4	4.7	21.5
Superior	28.7	2.1	21.5	16.9	8.0	22.8	16+	44.3	0.0	14.8	19.3	6.8	14.8
	33.3	0.7	20.0	23.0	5.2	17.8							
Age													
10-13	23.0	4.1	18.0	22.1	9.8	23.0	35-44	27.0	1.5	21.2	22.4	9.1	19.0
14-17	31.0	1.8	18.6	23.9	8.0	16.8	45-54	32.9	4.0	22.2	16.0	8.0	16.9
18-24	31.0	1.8	18.6	23.9	8.0	16.8	55-64	33.9	2.3	17.8	21.3	5.8	19.0
25-34	31.7	2.4	26.9	21.3	6.7	11.0	65+	27.6	2.6	23.0	17.8	7.2	21.7
	30.6	2.1	21.0	18.2	5.7	22.4							

TABLE 2 (continued)

Back-ground Variables	Arrangements					Back-ground Variables	Arrangements						
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV
Plate 4													
Total sample	31.1	11.4	22.7	1.0	0.7	33.1	Edu- cation						
Intelligence													
Subnor- mal	28.2	14.5	28.2	2.3	0.8	25.9	0- 6	30.8	13.3	24.3	2.6	1.3	27.9
Low Av.	31.4	13.8	25.9	0.8	0.0	28.0	7- 8	26.1	13.9	25.8	0.8	0.8	32.4
Average	32.6	9.6	22.6	1.2	0.9	33.0	9-11	34.8	8.8	19.9	1.0	0.3	35.2
High Av.	32.1	13.1	19.0	0.4	1.3	34.2	12	35.1	11.4	22.5	0.3	0.0	30.6
Superior	23.7	10.4	19.3	0.0	0.0	46.7	13-15	27.1	6.5	21.5	0.0	1.9	43.0
							16+	29.5	9.1	15.9	0.0	1.1	44.3
Age													
10-13	38.5	14.0	18.9	2.5	0.0	26.2	35-44	30.7	11.4	25.8	0.4	0.8	31.1
14-17	36.3	12.4	11.5	0.9	0.0	39.0	45-54	28.0	15.6	27.3	1.8	0.9	32.4
18-24	42.1	8.0	16.5	0.0	0.0	33.5	55-64	26.5	8.1	29.3	1.2	2.3	32.8
25-34	31.0	11.0	23.5	0.0	0.4	34.2	65+	20.4	9.9	29.6	2.6	1.3	36.2
Plate 5													
Total sample	35.8	21.2	12.4	13.9	7.4	9.3	Edu- cation						
Intelligence													
Subnor- mal	25.2	25.9	9.9	14.5	8.4	16.0	0- 6	30.1	20.4	11.0	14.3	9.7	14.6
Low Av.	35.1	14.2	11.7	15.9	9.6	13.4	7- 8	31.8	20.1	11.2	15.2	10.6	10.9
Average	36.2	21.3	12.1	13.6	8.0	8.8	9-11	40.9	20.3	13.2	12.8	5.4	7.4
High Av.	37.1	24.5	11.8	16.9	3.4	6.3	12	38.4	19.2	14.4	15.9	4.8	7.2
Superior	43.0	23.0	17.0	7.4	5.9	3.7	13-15	42.1	28.1	12.2	7.5	6.5	3.7
							16+	36.4	30.7	12.5	11.4	3.4	5.7
Age													
10-13	39.3	16.4	4.1	19.7	9.8	10.7	35-44	34.1	23.5	9.1	18.2	6.4	8.7
14-17	43.4	15.0	15.0	14.2	5.3	7.1	45-54	34.2	19.6	16.9	14.7	8.0	6.7
18-24	39.0	23.2	11.6	9.1	7.9	9.1	55-64	31.6	17.8	16.1	12.7	8.1	13.8
25-34	39.9	26.0	11.4	11.0	4.3	7.5	65+	26.3	21.1	14.5	13.2	11.8	13.2
Plate 6													
Total sample	33.9	43.0	1.5	7.1	8.4	6.1	Edu- cation						
Intelligence													
Subnor- mal	39.7	21.4	2.3	17.5	9.2	9.9	0- 6	41.8	29.5	2.3	11.7	8.4	6.5
Low Av.	38.5	33.5	2.9	8.4	8.8	7.9	7- 8	36.2	36.5	1.4	8.2	10.1	7.6
Average	33.1	43.6	1.3	7.0	8.6	6.1	9-11	34.1	45.0	2.0	7.1	7.4	4.4
High Av.	31.2	52.3	0.0	3.0	8.9	4.6	12	29.4	48.3	0.9	5.1	9.3	6.9
Superior	28.9	61.5	2.2	0.7	4.4	2.2	13-15	23.4	61.7	1.9	1.9	7.5	3.7
							16+	25.0	68.2	0.0	0.0	2.3	4.5
Age													
10-13	27.9	50.8	0.8	4.1	13.1	3.3	35-44	33.0	45.1	1.5	5.7	8.0	6.8
14-17	19.5	54.0	3.5	8.0	8.9	6.2	45-54	38.7	34.7	0.9	7.1	9.3	9.3
18-24	21.3	60.3	0.6	5.5	6.1	6.1	55-64	48.3	23.6	2.9	13.8	7.5	4.0
25-34	29.5	52.3	1.1	5.3	5.7	6.1	65+	48.7	24.3	2.0	7.9	11.8	5.3



TABLE 2 (continued)

Back-ground Variables	Arrangements					Back-ground Variables	Arrangements					
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO
Plate 7												
Total sample	3.6	7.4	6.6	36.8	44.6	1.1	Edu- cation					
Intelligence							0- 6	6.8	8.7	7.8	33.4	41.5
Subnor- mal	8.4	11.4	10.7	21.4	46.5	1.5	7- 8	5.2	9.5	7.9	32.9	43.2
Low Av.	6.3	8.8	5.0	36.0	42.3	1.7	9-11	2.0	6.4	5.4	38.2	47.0
Average	3.2	6.8	6.5	40.2	42.2	1.1	12	1.2	6.9	6.3	39.6	45.0
High Av.	1.3	7.6	6.8	37.6	46.0	0.8	13-15	1.9	4.7	4.7	42.1	46.8
Superior	0.0	3.7	5.9	32.6	57.8	0.0	16+	2.3	2.3	4.5	42.0	48.8
Age												
10-13	0.0	9.8	8.2	34.4	46.7	0.8	35-44	2.7	5.4	7.6	45.1	38.3
14-17	0.0	5.3	4.4	35.4	54.0	0.9	45-54	4.0	9.3	6.2	33.3	46.2
18-24	3.7	5.5	5.5	29.9	54.9	0.6	55-64	5.8	8.1	5.8	43.1	36.8
25-34	1.4	6.8	5.7	36.7	48.4	1.1	65+	11.2	9.9	9.9	30.9	35.5
Plate 8												
Total sample	7.9	8.1	1.5	57.6	22.6	2.4	Edu- cation					
Intelligence							0- 6	7.1	15.2	3.2	48.9	21.1
Subnor- mal	10.7	13.7	2.3	51.1	16.8	5.3	7- 8	10.1	9.2	1.6	52.2	24.8
Low Av.	10.0	10.9	2.5	50.2	23.4	2.9	9-11	7.8	7.1	1.0	61.2	21.6
Average	8.1	8.2	0.9	59.0	22.1	1.5	12	6.9	5.1	0.6	65.7	20.1
High Av.	6.8	4.2	1.3	62.9	22.4	2.5	13-15	6.5	0.9	0.9	60.8	27.1
Superior	2.2	3.0	3.0	60.0	28.2	3.7	16+	6.8	1.1	1.1	62.5	26.1
Age												
10-13	9.0	4.1	2.5	46.7	36.9	0.8	35-44	8.7	8.0	0.8	62.5	17.8
14-17	8.0	1.8	2.7	52.2	33.7	1.8	45-54	6.7	9.3	1.8	55.1	22.8
18-24	7.9	5.5	0.0	59.8	26.2	0.6	55-64	4.6	17.3	1.2	59.8	16.1
25-34	8.5	2.1	1.4	65.9	20.3	1.8	65+	9.9	17.8	3.3	45.4	17.8
Plate 9												
Total sample	14.7	5.9	30.9	16.5	20.0	11.9	Edu- cation					
Intelligence							0- 6	13.9	7.1	25.6	15.9	23.3
Subnor- mal	9.2	6.9	20.6	22.1	23.7	17.5	7- 8	12.2	6.3	25.3	18.8	19.9
Low Av.	15.1	8.8	24.7	16.7	21.8	13.0	9-11	13.9	5.4	34.5	18.6	19.3
Average	14.5	6.0	31.1	16.9	19.2	12.2	12	17.1	6.0	32.4	15.0	20.7
High Av.	18.6	3.0	34.2	15.2	21.5	7.6	13-15	17.8	2.8	41.1	15.9	14.0
Superior	14.8	5.2	44.5	9.6	14.8	11.1	16+	18.2	5.7	42.0	9.1	15.9
Age												
10-13	13.1	3.3	29.5	15.6	23.8	14.8	35-44	16.7	4.2	29.2	17.3	20.8
14-17	16.8	5.3	42.3	16.0	12.4	7.1	45-54	12.0	6.7	33.3	18.2	15.1
18-24	13.4	4.9	34.1	15.2	22.0	10.4	55-64	12.1	8.6	24.2	18.4	21.3
25-34	16.7	6.1	33.1	13.5	22.8	7.8	65+	16.5	8.6	22.4	16.5	19.1

TABLE 2 (continued)

Back-ground Variables	Arrangements					Back-ground Variables	Arrangements						
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV
Plate 10													
Total sample	65.4	6.7	8.5	13.9	2.8	2.7							
Intelligence							Edu- cation						
Subnor- mal	60.3	4.6	8.4	13.0	6.1	7.6	0- 6	59.6	5.8	7.8	16.2	5.5	5.2
Low Av.	65.7	7.1	7.5	12.6	4.6	2.5	7- 8	66.6	5.7	6.0	14.7	3.5	3.3
Average	66.4	7.6	8.0	13.4	2.3	2.3	9-11	66.2	6.8	9.5	13.9	1.7	2.0
High Av.	67.5	5.1	8.9	14.3	1.7	2.5	12	68.7	9.9	9.3	9.3	1.2	1.5
Superior	60.7	5.9	12.6	18.5	1.5	0.7	13-15	61.7	3.7	14.0	18.7	1.9	0.0
							16+	68.2	5.7	9.1	14.8	1.1	1.1
Age													
10-13	64.0	8.2	7.4	18.0	1.6	0.8	35-44	62.5	7.2	9.9	12.9	3.0	4.5
14-17	56.7	10.7	10.7	15.0	3.5	3.5	45-54	70.2	3.6	7.1	12.9	2.2	4.0
18-24	65.8	10.4	12.2	9.1	2.4	0.0	55-64	63.8	6.3	6.9	16.7	2.9	3.5
25-34	69.8	6.4	7.8	12.8	2.5	0.7	65+	63.8	3.9	7.2	16.5	4.6	3.9
Plate 11													
Total sample	8.7	24.5	15.7	6.2	14.3	30.7							
Intelligence							Edu- cation						
Subnor- mal	8.4	24.4	18.3	9.2	15.3	24.4	0- 6	7.8	24.9	15.6	8.7	13.3	29.8
Low Av.	12.6	25.1	8.8	9.6	14.2	29.7	7- 8	8.2	23.4	16.0	8.7	16.9	26.7
Average	8.2	25.3	15.6	5.9	14.4	30.6	9-11	11.8	27.4	13.5	4.7	12.2	30.4
High Av.	6.8	24.5	21.5	3.4	11.8	32.1	12	7.5	24.6	15.3	3.9	13.8	34.8
Superior	8.2	19.3	15.6	4.5	17.0	35.6	13-15	10.3	23.4	18.7	3.7	13.1	30.9
							16+	5.7	18.2	19.3	3.4	18.2	35.2
Age													
10-13	4.9	28.7	12.3	6.6	14.0	33.6	35-44	8.3	20.5	16.7	6.8	15.5	32.2
14-17	14.2	28.3	12.4	4.4	9.7	31.0	45-54	9.3	23.6	16.4	7.6	12.4	30.7
18-24	8.5	26.8	12.2	1.8	10.4	40.2	55-64	9.2	21.9	18.4	8.6	14.4	27.6
25-34	8.9	26.0	16.0	6.4	15.7	27.1	65+	6.6	24.3	17.8	5.9	19.7	25.7
Plate 12													
Total sample	1.7	14.1	51.8	6.2	4.1	22.1							
Intelligence							Edu- cation						
Subnor- mal	2.3	14.5	44.3	10.7	4.6	23.7	0- 6	1.9	13.6	49.2	7.1	4.5	23.7
Low Av.	2.1	16.7	49.0	5.4	3.8	23.0	7- 8	2.4	10.3	55.5	6.5	3.5	21.5
Average	1.7	12.5	55.2	6.0	3.6	20.9	9-11	0.3	11.8	55.1	5.4	3.7	23.7
High Av.	1.3	15.2	51.5	4.6	5.1	22.4	12	1.8	17.4	52.5	4.8	3.9	19.5
Superior	1.5	17.0	45.9	5.9	4.4	25.2	13-15	2.8	20.6	43.0	6.5	2.8	24.3
							16+	1.1	19.3	42.0	9.1	8.0	20.4
Age													
10-13	0.8	18.9	55.7	6.6	1.6	16.4	35-44	1.1	14.4	56.5	4.2	3.8	20.1
14-17	2.7	20.3	48.7	4.4	3.5	20.4	45-54	2.2	10.2	52.0	8.9	4.0	22.7
18-24	2.4	13.4	57.3	4.3	2.4	20.1	55-64	1.7	9.2	48.9	9.2	2.9	28.2
25-34	1.8	16.4	52.0	5.3	5.0	19.6	65+	1.3	13.8	40.1	6.6	7.9	30.3

TABLE 2 (continued)

Back-ground Variables	Arrangements					Back-ground Variables	Arrangements						
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV
Plate 13													
Total sample	14.7	52.1	4.7	2.1	10.2	16.1	Education						
Intelligence							0-6	23.0	30.8	8.1	2.9	14.3	21.1
Subnormal	19.1	32.8	13.0	4.6	15.3	15.3	7-8	15.8	45.7	6.3	4.4	9.8	18.0
Low Av.	18.8	45.2	5.0	2.9	13.0	15.1	9-11	12.5	60.2	4.1	1.4	8.1	13.9
Average	14.1	54.5	4.4	1.9	9.2	15.8	12	10.2	61.2	2.7	0.9	10.5	14.4
High Av.	14.3	53.2	3.4	1.7	8.9	18.6	13-15	12.2	68.3	0.9	0.0	6.5	12.2
Superior	8.2	67.4	0.7	0.7	7.4	15.6	16+	9.1	71.6	1.1	0.0	8.0	10.2
Age													
10-13	29.5	31.1	4.1	1.6	8.2	25.4	35-44	12.1	59.1	3.8	2.3	8.3	14.4
14-17	14.2	54.9	5.3	1.8	8.0	16.8	45-54	12.4	47.6	7.6	3.6	12.9	16.0
18-24	11.6	62.8	2.4	0.6	8.0	14.6	55-64	17.3	41.4	6.9	2.3	13.8	18.4
25-34	10.3	65.5	2.8	0.7	8.2	12.5	65+	20.4	38.2	5.9	4.6	14.5	16.5
Plate 14													
Total sample	69.1	4.7	1.2	4.3	2.1	18.7	Education						
Intelligence							0-6	54.1	8.1	2.3	7.1	2.6	25.9
Subnormal	49.6	9.9	3.1	9.2	0.8	27.5	7-8	62.6	6.8	1.4	5.7	3.3	20.1
Low Av.	60.7	7.9	1.7	4.2	5.0	20.5	9-11	76.7	2.7	0.3	3.0	2.0	15.2
Average	70.2	4.1	1.2	4.8	2.0	17.6	12	76.2	2.4	0.9	2.4	0.9	17.1
High Av.	78.1	1.7	0.4	1.3	1.3	17.3	13-15	79.5	2.8	0.9	1.9	0.9	14.0
Superior	81.5	2.2	0.0	1.5	0.7	14.1	16+	82.9	1.1	1.1	2.3	2.3	10.2
Age													
10-13	60.7	2.5	0.8	2.5	3.3	30.3	35-44	73.1	5.7	0.4	2.7	1.1	17.0
14-17	78.7	1.8	0.9	1.8	0.0	16.8	45-54	67.6	5.3	2.2	3.6	1.8	19.6
18-24	83.0	1.2	0.0	2.4	3.0	10.4	55-64	62.1	6.9	3.5	6.9	1.2	19.6
25-34	72.3	3.2	0.7	5.3	2.1	16.4	65+	52.0	9.9	1.3	7.9	5.3	23.7
Plate 15													
Total sample	28.3	8.5	10.7	14.3	2.5	35.8	Education						
Intelligence							0-6	26.2	11.0	13.6	13.6	2.3	33.4
Subnormal	28.2	12.2	13.7	16.0	0.8	29.0	7-8	31.6	7.1	14.1	13.6	3.3	30.2
Low Av.	23.0	8.4	10.0	15.5	3.3	39.7	9-11	27.0	9.5	9.5	13.2	2.0	38.9
Average	27.9	8.4	10.9	13.3	2.8	36.6	12	26.4	9.0	8.1	15.0	3.0	38.4
High Av.	31.7	8.4	11.4	12.7	2.1	33.8	13-15	33.7	3.7	2.8	16.8	0.9	42.1
Superior	34.1	5.9	5.9	17.8	2.2	34.1	16+	27.3	5.7	9.1	17.0	2.3	38.6
Age													
10-13	36.9	7.4	7.4	9.8	4.9	33.6	35-44	29.5	6.4	12.9	13.3	3.0	34.8
14-17	29.2	7.1	6.2	16.0	2.7	39.0	45-54	33.3	8.9	9.8	11.6	3.1	33.3
18-24	22.0	7.3	8.5	20.7	3.7	37.8	55-64	24.2	9.8	17.3	8.1	0.6	40.3
25-34	26.0	8.2	7.5	19.6	1.4	37.4	65+	27.0	13.8	14.5	13.2	2.0	29.6

TABLE 2 (continued)

Back-ground Vari- ables	Arrangements					Back-ground Vari- ables	Arrangements							
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV	
Plate 16														
Total sample	0.7	86.4	2.5	2.1	3.5	4.9	Education							
Intelligence														
Subnormal								0- 6	1.6	74.8	4.2	2.3	8.1	9.1
Low Av.	1.5	72.5	4.6	3.8	8.4	9.2		7- 8	0.3	84.9	3.3	2.4	4.6	4.4
Average	1.3	79.9	2.5	3.8	5.9	6.7		9-11	0.3	88.6	2.4	2.7	1.0	5.1
High Av.	0.4	88.4	2.8	1.6	2.4	4.3	12	0.6	93.0	0.3	1.2	1.2	3.6	
Superior	0.8	90.7	0.4	1.7	3.0	3.4	13-15	0.9	92.6	0.9	1.9	1.9	1.9	
	0.0	91.1	2.2	1.5	2.2	3.0	16+	0.0	92.0	3.4	2.3	2.3	0.0	
Age														
10-13	0.0	88.5	1.6	2.5	2.5	4.9	35-44	0.8	90.5	2.3	2.3	1.1	3.0	
14-17	0.0	82.3	3.5	2.7	2.7	8.9	45-54	0.4	85.3	1.3	1.3	4.9	6.7	
18-24	0.0	89.6	0.6	1.2	4.3	4.3	55-64	0.6	84.0	2.9	2.9	5.2	4.6	
25-34	1.1	90.8	1.8	2.1	2.1	2.1	65+	2.0	73.0	7.2	2.6	7.2	7.9	
Plate 17														
Total sample	33.1	18.9	12.2	22.5	10.6	2.7	Education							
Intelligence														
Subnormal								0- 6	27.9	19.8	16.5	21.7	11.7	2.6
Low Av.	29.8	17.5	14.5	22.1	11.4	4.6		7- 8	29.9	17.1	11.4	23.7	12.8	4.9
Average	24.3	20.9	17.2	24.3	10.0	3.3		9-11	34.8	16.9	11.5	24.0	10.5	2.4
High Av.	33.6	18.0	12.1	22.9	10.6	2.7	12	35.1	18.9	12.9	23.4	8.4	1.2	
Superior	37.1	18.6	9.7	20.7	11.8	2.1	13-15	41.1	22.4	7.5	18.7	8.4	1.9	
	41.5	21.5	6.7	22.2	7.4	0.7	16+	42.0	25.0	5.7	17.0	9.1	1.1	
Age														
10-13	33.6	19.7	11.5	25.4	8.2	1.6	35-44	34.8	20.1	10.6	22.0	10.2	2.3	
14-17	35.4	14.2	10.6	29.2	7.1	3.5	45-54	28.0	16.4	16.0	22.7	13.8	3.1	
18-24	45.1	20.7	6.1	17.7	9.1	1.2	55-64	28.8	21.9	13.2	20.7	12.1	3.5	
25-34	33.1	21.4	10.7	22.8	9.6	2.5	65+	27.0	12.5	19.7	23.7	13.2	3.9	
Plate 18														
Total sample	38.0	7.3	9.1	23.6	9.5	12.5	Education							
Intelligence														
Subnormal								0- 6	35.0	7.8	4.9	26.9	10.7	14.9
Low Av.	24.4	8.4	8.4	23.7	17.5	17.5		7- 8	34.5	4.9	7.9	28.0	11.2	13.3
Average	39.7	5.9	6.7	25.1	11.7	10.9		9-11	38.9	6.1	10.8	24.0	8.1	12.2
High Av.	38.6	7.6	7.7	25.5	7.8	12.6	12	39.3	10.2	8.7	20.7	9.3	11.7	
Superior	40.9	7.6	12.2	19.0	9.7	10.6	13-15	49.6	7.5	18.7	12.2	4.7	7.5	
	40.0	7.4	17.0	17.8	6.7	11.1	16+	39.8	9.1	13.6	17.0	10.2	10.2	
Age														
10-13	36.9	2.5	5.7	32.8	9.8	12.3	35-44	40.6	9.5	10.2	20.5	8.7	10.6	
14-17	40.0	5.3	11.5	17.7	10.6	15.0	45-54	35.6	6.7	9.8	27.6	11.1	9.3	
18-24	40.2	8.5	12.2	17.7	8.5	12.8	55-64	38.5	9.8	6.3	24.7	6.9	13.8	
25-34	37.7	8.5	11.0	20.3	10.0	12.5	65+	34.2	3.9	3.9	31.6	10.5	15.8	



TABLE 2 (continued)

Back-ground Variables	Arrangements					Back-ground Variables	Arrangements						
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV
Plate 19													
Total sample	4.3	17.1	27.5	1.7	4.1	44.9	Edu- cation						
Intelligence							0- 6	5.2	23.0	29.2	2.9	7.8	32.1
Subnor- mal	6.1	25.2	22.9	4.6	7.6	33.6	7- 8	6.8	16.6	29.4	2.5	5.2	39.4
Low Av.	4.6	20.1	27.2	1.7	9.2	37.2	9-11	3.7	16.2	27.7	1.0	3.4	48.0
Average	4.7	16.5	29.1	1.5	3.3	44.8	12	3.0	15.0	28.2	0.3	1.8	51.6
High Av.	2.1	14.8	29.5	0.8	1.3	51.5	13-15	1.9	15.9	23.4	2.8	1.9	54.2
Superior	3.7	11.1	24.5	1.5	0.7	58.5	16+	1.1	11.4	22.7	0.0	0.0	64.8
Age													
10-13	4.1	18.0	28.7	4.1	2.5	42.6	35-44	3.8	19.3	27.3	1.5	4.2	44.0
14-17	3.5	14.2	29.2	0.0	4.4	48.7	45-54	7.1	16.4	26.7	1.3	4.9	43.6
18-24	3.0	9.1	28.0	1.8	1.2	56.7	55-64	4.6	16.7	36.8	1.7	5.2	35.1
25-34	3.2	19.2	21.7	2.1	2.1	51.6	65+	5.3	20.4	30.3	0.7	9.2	34.2
Plate 20													
Total sample	11.3	2.3	3.1	36.2	44.6	2.5	Edu- cation						
Intelligence							0- 6	14.9	3.6	3.9	36.9	37.6	3.2
Subnor- mal	16.0	3.1	7.6	38.2	31.3	3.8	7- 8	15.0	2.5	3.5	35.9	40.3	2.7
Low Av.	14.2	3.3	2.1	36.4	41.0	2.9	9-11	10.5	2.4	1.7	35.5	48.3	1.7
Average	12.0	2.3	2.3	35.6	45.6	2.1	12	7.5	1.8	3.6	37.5	47.4	2.1
High Av.	7.2	0.8	4.2	35.8	50.2	1.7	13-15	5.6	0.0	2.8	36.5	54.2	0.9
Superior	5.9	2.2	3.0	36.3	48.9	3.7	16+	8.0	2.3	2.3	30.7	52.3	4.5
Age													
10-13	7.4	0.0	5.7	41.0	43.4	2.5	35-44	10.6	1.9	2.7	37.5	46.6	0.8
14-17	9.7	1.8	0.0	42.3	43.4	2.7	45-54	16.0	2.7	4.4	33.8	40.9	2.2
18-24	7.9	1.8	0.6	39.6	47.6	2.4	55-64	13.8	2.3	5.2	39.1	36.8	2.9
25-34	9.6	1.8	2.8	29.9	53.0	2.8	65+	14.5	6.6	2.6	32.2	39.5	4.6
Plate 21													
Total sample	1.7	84.3	1.3	0.9	3.6	8.3	Edu- cation						
Intelligence							0- 6	3.9	74.2	3.2	2.6	7.1	9.1
Subnor- mal	3.8	68.7	6.1	3.1	6.9	11.4	7- 8	2.7	82.7	1.6	1.4	5.7	5.7
Low Av.	5.0	78.2	2.1	2.1	2.5	10.0	9-11	0.0	85.5	1.0	0.0	2.0	11.5
Average	0.9	87.0	0.8	0.5	4.3	6.4	12	0.6	92.1	0.0	0.0	0.3	6.9
High Av.	0.8	88.6	0.0	0.0	3.0	7.6	13-15	1.9	82.3	0.0	0.0	3.7	12.2
Superior	0.0	87.4	0.0	0.0	0.0	12.6	16+	0.0	94.3	0.0	0.0	0.0	5.7
Age													
10-13	1.6	84.4	0.8	2.5	3.3	7.4	35-44	1.9	84.9	1.9	0.8	1.5	9.1
14-17	1.8	85.6	0.0	1.8	1.8	8.9	45-54	2.2	84.4	1.3	0.9	3.6	7.6
18-24	0.6	87.2	0.6	0.0	2.4	9.1	55-64	2.3	81.1	1.7	0.6	9.2	5.2
25-34	0.7	88.3	0.0	0.4	1.1	9.6	65+	3.3	75.0	3.9	1.3	8.6	7.9

TABLE 2 (continued)

Back-ground Vari- ables	Arrangements						Back-ground Vari- ables	Arrangements					
	VOL	VLO	OLV	OVL	LVO	LOV		VOL	VLO	OLV	OVL	LVO	LOV
Plate 22													
Total sample	0.9	6.9	21.0	51.1	19.5	0.6							
Intelligence							Edu- cation						
Subnor- mal	3.1	6.9	22.1	47.3	19.8	0.8	0- 6	2.3	6.2	20.7	48.9	20.4	1.6
Low Av.	0.4	6.7	20.1	51.9	20.5	0.4	7- 8	0.3	6.0	22.9	54.1	16.0	0.5
Average	1.1	7.2	21.3	51.1	18.4	0.9	9-11	2.0	5.4	18.6	50.0	23.3	0.7
High Av.	0.4	6.3	18.1	51.1	24.1	0.0	12	0.0	9.3	21.3	49.5	19.8	0.0
Superior	0.0	5.2	25.9	51.9	17.0	0.0	13-15	0.0	11.2	21.5	48.6	18.7	0.0
							16+	0.0	3.4	20.4	57.9	18.2	0.0
Age													
10-13	0.8	3.3	12.3	62.3	21.3	0.0	35-44	0.8	3.8	26.1	51.1	17.4	0.8
14-17	0.9	4.4	16.8	48.7	28.3	0.9	45-54	1.3	9.8	15.6	53.3	15.1	0.9
18-24	1.2	9.8	18.9	42.1	28.0	0.0	55-64	1.2	8.6	24.2	48.3	17.3	0.6
25-34	0.7	7.1	19.2	53.4	19.2	0.4	65+	0.7	6.6	27.0	48.0	16.5	1.3
Plate 23													
Total sample	30.2	22.9	18.9	2.8	2.7	22.6							
Intelligence							Edu- cation						
Subnor- mal	25.9	23.7	19.1	3.8	3.8	23.7	0- 6	32.4	22.4	17.2	3.9	3.9	20.4
Low Av.	30.5	23.8	20.9	2.9	4.6	17.2	7- 8	31.6	22.6	17.4	3.3	3.0	22.0
Average	31.9	20.9	17.4	2.8	2.5	24.3	9-11	31.4	22.0	21.0	2.4	1.7	21.6
High Av.	29.1	24.1	21.9	3.0	2.1	19.8	12	30.0	22.5	19.2	2.1	2.7	23.4
Superior	25.2	23.2	18.5	1.5	0.0	26.7	13-15	23.4	25.2	20.6	1.9	1.9	27.1
							16+	21.6	27.3	20.4	2.3	1.1	27.3
Age													
10-13	39.3	22.1	18.9	1.6	3.3	14.8	35-44	32.2	23.9	17.8	3.8	3.0	19.3
14-17	35.4	15.0	16.9	1.8	1.8	29.2	45-54	26.2	24.4	20.0	5.3	1.3	22.7
18-24	25.0	20.1	21.3	1.8	0.6	31.1	55-64	28.8	22.4	20.7	2.9	4.0	21.3
25-34	26.7	25.6	20.3	1.4	3.2	22.8	65+	34.2	23.7	13.8	2.6	3.9	21.7
Plate 24													
Total sample	20.4	40.0	4.9	4.7	11.9	18.1							
Intelligence							Edu- cation						
Subnor- mal	16.8	36.6	11.4	6.1	6.9	22.1	0- 6	19.4	35.0	6.2	3.9	10.7	24.9
Low Av.	26.4	32.2	5.0	3.3	9.2	23.8	7- 8	21.8	35.9	7.3	5.7	10.3	18.8
Average	19.6	40.7	4.3	4.5	14.4	16.5	9-11	22.6	40.6	4.4	5.4	13.9	13.2
High Av.	19.4	41.8	5.1	5.9	10.1	17.7	12	21.0	43.2	3.6	3.6	12.0	16.5
Superior	20.0	48.9	2.2	4.4	11.9	12.6	13-15	13.1	46.8	2.8	3.7	15.0	18.7
							16+	17.0	52.3	0.0	5.7	12.5	12.5
Age													
10-13	23.0	32.8	3.3	4.9	18.9	17.2	35-44	22.0	39.0	6.1	4.2	11.0	17.8
14-17	20.3	36.3	3.5	8.9	17.7	13.3	45-54	19.6	40.9	7.6	1.3	11.1	19.6
18-24	21.3	41.5	6.1	5.5	15.9	9.8	55-64	16.7	37.4	4.6	5.8	8.6	27.0
25-34	20.0	51.6	2.1	3.6	10.3	12.5	65+	21.1	28.2	5.9	7.2	7.9	29.6

TABLE 2 (continued)

TABLE 2 (continued)													
Back-ground Variables	Arrangements					Back-ground Variables	Arrangements						
	VOL	VLO	OLV	OVL	LVO		LOV	VOL	VLO	OLV	OVL	LVO	LOV
Plate 25													
Total sample	31.2	18.6	13.2	25.3	4.6	6.5	Education						
Intelligence							0-6	30.1	20.7	11.3	25.6	6.2	6.2
Subnormal	27.5	19.1	9.2	28.2	8.4	7.6	7-8	26.1	20.4	11.7	29.4	4.9	7.3
Low Av.	27.6	22.6	11.3	25.5	4.6	8.4	9-11	29.1	15.5	15.5	29.1	3.7	7.1
Average	31.1	17.6	14.0	25.8	5.5	6.0	12	35.1	17.7	14.4	23.4	5.1	4.2
High Av.	35.9	19.8	14.3	24.1	1.3	4.6	13-15	36.5	17.8	18.7	17.8	1.9	7.5
Superior	33.3	14.8	22.2	20.7	1.5	7.4	16+	42.0	18.2	18.2	10.2	2.3	9.1
Age													
10-13	28.7	11.5	19.7	28.7	4.1	7.4	35-44	36.4	18.2	13.3	22.7	4.5	4.9
14-17	29.2	16.9	21.2	22.1	3.5	7.1	45-54	26.2	20.9	14.2	27.1	4.4	7.1
18-24	28.0	17.7	20.7	19.5	6.1	7.9	55-64	31.1	22.4	5.2	30.5	5.2	5.8
25-34	36.0	17.4	13.9	22.1	3.6	7.1	65+	27.6	22.4	7.2	32.9	5.3	4.6

## C. RESULTS

In Table 2 we have presented the frequency with which each of the six possible arrangements for each plate of the *PAT* were selected in our sample of 1,500 cases, along with breakdowns by intelligence, education, and age.<sup>3</sup> The arrangements have been indicated by means of a code according to which *V* refers to a triangle, *O* to a circle, and *L* to a rectangle.

## D. SUMMARY

The Tomkins-Horn Picture Arrangement Test has been administered to a representative sample of the United States population over nine years of age. Information on education, sex, marital status, religion, age, occupation, occupation of family head, race, class, rural-urban residence, geographical area, and intelligence was obtained for each subject. Of these background variables only age, education, and intelligence show a marked effect on *PAT* performance and have therefore been employed in the computation of percentage frequencies for the six possible arrangements on each plate of the test.

<sup>3</sup>These three variables show the greatest effect on *PAT* performance and have therefore been selected for presentation here. Additional normative data based on breakdowns by sex, race, marital status, religion, occupation, geographical area, rural-urban residence, and occupation of family head will be published as part of the *PAT* manual at a later date.

## REFERENCES

1. CENTERS, R. *The Psychology of Social Classes*. Princeton: Princeton Univ. Press, 1949.
2. THORNDIKE, R. L. Two screening tests of verbal intelligence. *J. Appl. Psychol.*, 1942, **26**, 128-35.
3. THORNDIKE, R. L., & GALLUP, C. H. Verbal intelligence of the American adult. *J. Gen. Psychol.*, 1944, **30**, 75-85.
4. TOMKINS, S. S. The Tomkins-Horn Picture Arrangement Test. *Trans. N. Y. Acad. Sci.*, 1952, **15**, 46-50.
5. WECHSLER, D. *The Measurement of Adult Intelligence*. Baltimore: Williams & Wilkins, 1944.

*Department of Psychology*  
*Princeton University*  
*Princeton, New Jersey*



## THE RATIONALE, CONSTRUCTION, AND PRELIMINARY TRY-OUT OF THE SYNONYM VOCABULARY TEST\*

*Department of Psychology, Pennsylvania State University*

---

WILLIAM M. LEPLEY

---

### A. INTRODUCTION AND RATIONALE

The instrument to be described had its origins in general theory concerning behavior variability. This theoretical development is contained in a previous publication of the writer (2). The principal thesis of this development is embodied in the following quotation from the reference cited.

It is biologically important that the behavior of organisms be unpredictable; behavior variability has survival value; behavior variability is an aspect of and a contribution to, adaptability. This proposition seems to be especially applicable in situations wherein organisms stand in the relation of antagonists; the hunter and the hunted, the oppressor and the oppressed, paired debators, paired fencers, paired chess players. Variability of behavior is a mode of attack and a mode of defense.

The writer then invoked the Hullian Postulate No. 8 (1) (which has to do with the reactive inhibition principle) in a modified form, as the principle underlying the generation of behavior variability. Finally, in the publication cited, the writer noted: "... the writer calls attention to the long-accepted finding that the vocabulary test score is one of the better single indices of intelligence. The large vocabulary is interpreted here as a richly variable behavior repertoire." In a footnote to this last he stated: "The writer has under way a project designed to test the hypothesis that the large synonym vocabulary is an especially valid indication of high intelligence." The project referred to in the immediately preceding quotation is the one to be described in the present paper.

A brief, approximate statement of this rationale follows. Though two individuals may be essentially equal in concept repertoire, they may differ significantly in language behavior variability, in the synonym vocabularies, by means of which their concepts are expressed. It is proposed that such differences, if they exist, might reflect differences in adaptability or intelligence, as operationally defined by way of some criteria yet to be identified;

---

\*Received in the Editorial Office on October 16, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

and might do so with refinements of precision and meaning not yet achieved by other techniques for vocabulary assessment.

### B. THE PROCESS OF CONSTRUCTION

The initial plan for the instrument called for the construction of an unspecified number of matching-type items, exemplified in Table 1 which is the short, practice item used with the instructions to the examinee, in the present form of the test.

TABLE 1

A. THIN	B. LEAVE	C. DESTROY	D. PROBLEM	E. ABLE
1. desert		6. abandon		11. resign
2. wreck		7. expert		12. ruin
3. puzzle		8. abolish		13. mystery
4. lean		9. riddle		14. slim
5. slender		10. clever		15. skilled

As will be explained, the actual test items, in the above format, have 30 response words per item rather than 15 as shown for the practice item.

### C. ESTABLISHMENT OF THE BASIC FILE

The basic file consists of 502 cards. The heading of each card is a frequently used word taken from the Thorndike-Lorge (4) count. On each card are entered synonyms and analogous words, related to the heading word, varying greatly in frequency of use according to this same count. The construction steps are described in the following paragraphs.

Each "AA" word in the Thorndike-Lorge *G*-count<sup>1</sup> (4) was sought in Webster's Dictionary of Synonyms (5). Six hundred and fourteen (of the total of 1,069) "AA" words were found and entered, in an alphabetical, card file.

All synonyms and analogous words for each of the 614 words were transcribed on the appropriate cards in this file. In this process, the synonym and analogous word entries for multi-meaning words were segregated by meaning, or sense, category.<sup>2</sup>

All of the entries yielded by the synonym dictionary were then sought in the Thorndike-Lorge, *G*-count and the frequency for each was recorded in

<sup>1</sup>The *G*-count is the general count, "AA" words are those that were counted 100 or more times per million words. "A" words are those counted at least 50 times per million but not so many as 100 per million.

<sup>2</sup>It may be appropriate to mention that for each "AA" word card, within each meaning category, the synonym and analogous word entries were arranged in a descending order of frequency; this for later convenience.

the 614-card file.<sup>3</sup> Those not found were assigned a frequency of zero.

A separate card was made for each synonym and analogous word entry in the original 614-word file. The cards in this second file showed the word, its frequency in the *G*-count, and entries denoting the one or more "AA" words under which it had appeared. This process generated a file of 6,380 cards with a total of 20,646 word entries.

The original file of 614 cards was sorted in a search for all "AA" words having frequency counts as low as five per million among the synonyms in at least one meaning or sense category. Those which did not meet this criterion were removed from the file. This sort left 502 "AA" word items in the file. (From this point onward, this 502-item alphabetized deck will be referred to as "the basic file.")

#### D. THE SELECTION OF THE STEM WORDS

The 80 stem words were chosen from the basic file in the following manner.

The basic file was divided into 80 approximately equal segments. This number was chosen since it was proposed that the final form of the instrument would contain 80 stem words, divided into 16 matching-type items; each with five stem words and 30 response words chosen from among the associated synonyms and analogous words.<sup>4</sup> The item format is shown by example in an earlier paragraph.

The 80 equal segments then were searched card by card for the best word in each segment. The criteria for selection in this process were: (a) large number of synonyms and analogous word entries within some meaning category; (b) wide dispersion of frequency count values for the synonym and analogous word entries, the ideal in this being from "A" to zero; (c) freedom of this dispersion from an excess of words of such form as to clearly suggest association with the stem word. For example, for the stem word *ADD*, the synonym *superadd* occurs with a very low frequency count. Such words as the latter were not used.

The first selection of the 80 stem words, according to these criteria, was

<sup>3</sup>The next step was to search Webster's Unabridged Dictionary (6) for synonyms and related words in meaning and usage categories not represented in the synonym dictionary; this for each of the 614 words in the original file. The words yielded by this procedure were added to the original file with their *G*-count frequencies. Later, it was decided that, in order to keep the process precisely describable, the gleanings from the unabridged dictionary would not be used.

<sup>4</sup>Within this type of matching item, each of the 30 response words served as a correct match for one stem word and as a mislead for each of the other four stem words.

accomplished by one judge, the author. The sort was then repeated with the aid of an additional judge. In the course of this second sort, the two judges attempted to eliminate as much of the semantic overlap among the 80 stem words as they could detect with this informal, conference procedure.

#### E. SELECTION OF THE RESPONSE WORDS

This part of the process consisted of the author's going through the deck of 80 and tentatively selecting the response words. Since it was intended that each stem word would be represented by approximately six correct matches, in each of the 16 matching-type items, the ideal number of selections for each stem word was six. The actual number of selections varied from four to eight. The details of procedure were as follows: (a) Wherever possible, the six (more or less) selections were made from what was judged to be the most generally familiar meaning category offered by the synonym dictionary. This was not always possible, because such categories did not always meet the criteria noted in the foregoing paragraphs. That is to say, the most familiar meaning category (as judged by the author) sometimes did not provide an adequate number of entries. In such cases the selections were made within the most familiar meaning category that did fulfill these requirements. (b) As before stated and implied, the ideal meaning category contained at least six entries, ranging in *G*-count from "A" to zero and with the entries evenly distributed over this range. Therefore the six (more or less) tentatively identified response words were selected to conform as closely as possible to this number, range and distribution.

This next step was designed for two purposes: (a) to insure that no response word was used more than once among the 480 response words and, (b) to remove more of the semantic overlap mentioned above.<sup>5</sup> The tentatively selected response words on each of the 80 stem word file cards were sorted against the 6,380-word file of all synonyms and analogous words noted in a foregoing paragraph. As implied above, this sorting removed all response word overlap in the 80-item file of stem words. Wherever overlap occurred, substitutions were made according to principles detailed above. The processes described were then applied to the substitute.

---

<sup>5</sup>It should be confessed that the author, at this time, was not confident that the several processes used got rid of all semantic overlap, even within the sub-decks of five. He ventures a guess that a complete sort of the English language, or any other language, accomplishing the elimination of all semantic overlap would leave a small number of residual concepts closely resembling the universals long of interest to philosophers.



## F. TRY OUT FOR ADEQUACY OF FORMAT AND INSTRUCTIONS

Before actually assembling and reproducing Form 1 (Experimental) a short form consisting of four items of the type and length described was put together. These items were constructed hastily with leftover materials. Instructions were prepared with a degree of simplicity designed to make them usable below the college level. This form was reproduced and administered to 98 undergraduates who were invited to comment and complain if appropriate. On the basis of this try out, minor changes were made in the instructions but not in the format. These test papers were scored for odd and even item scores and the two series correlated by means of the product-moment formula. The  $r$  yielded was .73. According to the Spearman-Brown prophecy formula, this value corrected to .84, the whole-test reliability estimate for this short form.

## G. TRY OUT OF FORM I (EXPERIMENTAL)

Form I (Experimental) was assembled with revised instructions, reproduced by mimeographing, and administered to 306 students with no time limit. The time required by these students to complete the test varied from 26 minutes to 75 minutes. Approximately 80 per cent completed the test within 50 minutes.

The students in this sample were distributed according to academic status as follows: 81 freshmen, 149 sophomores, 52 juniors, 20 seniors, three graduate students, and one special student. Table 2 shows the population characteristics of the whole group; and the distribution characteristics for the whole population and for those sub-populations of appreciable size. Attention is invited to the means of the raw scores, which exhibit an increase from the freshman group through the junior group. No attempt is made to interpret

TABLE 2  
COMPOSITION AND DESCRIPTION OF THE TRY OUT GROUP\*

	Freshmen	Sophomores	Juniors	Seniors	Graduates	Special Students	All
N	81	149	52	20**	3**	1**	306
M	283.2	295.7	319.8	—	—	—	297.5
SD	40.1	46.4	46.4	—	—	—	46.3
Mdn	—	—	—	—	—	—	299
Mode	—	—	—	—	—	—	294

\*All computations based on raw scores (number right) and ungrouped data.

\*\*Computations not made for small N's.

acteristics of the whole group; and the Synonym Vocabulary Test, raw score (number of correct responses) distribution characteristics for the whole population and for those sub-populations of appreciable size. Attention is invited to the means of the raw scores, which exhibit an increase from the freshman group through the junior group. No attempt is made to interpret

these apparent differences. They are likely attributable to two sets of factors, one selective and the other vocabulary increment.

#### H. RELIABILITY OF FORM I (EXPERIMENTAL)

The principal purpose of the test administration described above was to determine the instrument's reliability characteristics and to improve them if necessary. The 306 papers were scored for odd and even items and the product-moment correlation between the two arrays was computed. The results of this operation are shown in Table 3.

TABLE 3  
SPLIT-HALF RELIABILITY\*  
N = 306

	Odd scores	Even scores
M	150.9	146.6
SD	23.1	23.9
$r_{11}$		.94
$r_{211}$		.97

\*All computations based on raw scores (number right) and ungrouped data.

The estimated reliability for the whole test, while appreciable, was judged to be improvable. The steps taken in an attempt to improve it are described in a later section.

#### I. CLUES TO VALIDITY

The findings submitted here are not to be considered as the results of carefully planned validation studies. The impatient author was unable to withstand the temptation to use incidentally available data in these minor studies. They may afford some faint hints about the nature of the variable being measured.

##### 1. *Correlation With Final Examination Grades*

From the total population of 306 cases, 202 cases were segregated for whom raw scores from a 70-item, final examination in elementary psychology were available. These cases were from eight sections taught by six different instructors. For this population, the corrected split-half reliability of the 70-item examination was computed to be .80. The coefficient of correlation between Synonym Vocabulary Test scores and the final examination scores was computed to be .32 with a standard error of .06.

## 2. *Correlation With Course Grades*

The present writer was the instructor responsible for two of the eight sections mentioned above. The populations for these two sections were isolated, combined, and the computations described above were applied to the data from this group of 93 students.

The second variable, here included the 70-item final examination scores plus the scores for 314 additional objective items used over the span of the semester. The resulting coefficient was .52 with a standard error of .08.

These meager clues to validity are left to the reader's interpretation.

## J. ITEM ANALYSIS AND REVISION

Since it is well known that reliability indices shrink when measures are applied to restricted ranges, it was judged necessary to refine the instrument. Accordingly the following steps were executed.

The total population ( $N = 306$ ) was divided at the median and a separate item response count was made for the high and low scoring populations. The data yielded by this process were used as an adjunct to judgment in an inspection of the rational key. Four judges<sup>6</sup> working as a group, participated in this process. Wherever the item response count appeared to be out of line with the rational key, the item was carefully inspected and amended if judged to be faulty. This combination of statistics with judgment may need some defense, or at least, clarification. Examples should help. Item XIII in the experimental form was furnished with the stem words: PAY, LAW, DECLARE, HARD, and CHANGE. Among the response and mislead words was the word "warrant," which the synonym dictionary listed as a synonym for DECLARE. The inspection, judgment, and sorting processes, described earlier, failed to call attention to PAY and LAW as possibly reasonable and defensible responses. This and similar items were amended by replacing such response words with others (judged free of such fault) and matched as closely as possible on frequency of use. In other instances the judges in this latter process were not able to account for popular wrong responses. For these a special count was made in the highest scoring 10 per cent. If the wrong response remained popular in this population the word was discarded and replaced. If it did not, the item was left unamended with an unfounded hope that the response count would fall into line with the rational key when the instrument was used at higher levels along the continuum of the variable being

---

<sup>6</sup>The judges mentioned here, and earlier, included Dr. and Mrs. Richard P. Barthol and Mrs. William M. Lepley.

measured. It should be also stated that, because of the great range of difficulty deliberately built into this instrument, many obviously nondiscriminating response words were revealed in this analysis. These words were not replaced because the instrument was constructed to be usable over a wide range of trait variation. According to this complex process of counting and judging, which defies precise description in detail, 15 response words were replaced. These were distributed through 15 of the 16 matching-type items with no more than two substitutions in any item. It is judged that the principal effect of this process was to refine out still more of the semantic overlap sought by other procedures described in foregoing sections. It is hoped that the reliability will be revealed as improved.

#### K. CONSTRUCTION OF FORM I

The construction of this form consisted principally of re-arranging the order of the 16 matching-type items of Form I (Experimental) and dividing the new arrangement into Part *A* and Part *B*, each consisting of eight items. This was done with the hope that the two parts might be useful as adequately reliable, equivalent forms. Further, the format was altered slightly and modified instructions were written to permit the use of answer sheets and machine scoring.

#### L. ITEM REARRANGEMENT

The total *N* of 306 was used to compute mean scores and standard deviations for the 16 items. Table 4 shows these values. As a first step the items were arranged in an order of difficulty based on these mean error scores, beginning with the least difficult items. The items as in this rank order were then sorted into two categories by means of a scheme, which, under other circumstances, might be described as an ABBA schedule. Specifically, the item in Rank 1 was sorted into Category *A* (later to correspond with Part *A*); the items in Ranks 2 and 3 were sorted into Category *B*. Items in Ranks 4 and 5 were sorted into Category *A* and so on throughout the 16 ranks. Further adjustments were made so that, as finally sorted the mean total error scores for the two categories differed by only 0.2. The order of difficulty for each category thus was arranged to place the least difficult items first, and Categories *A* and *B* became Parts *A* and *B* in this revision. According to these procedures Form I of the Synonym Vocabulary Test was assembled.

#### M. GENERAL REMARKS

The foregoing step by step description has been kept uncluttered by remarks general in nature but possibly of some significance. The writer wishes



TABLE 4  
MEANS AND STANDARD DEVIATIONS FOR THE 16 MATCHING TYPE ITEMS\*  
 $N = 306$

	Item numbers															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M	12.7	11.5	9.8	15.4	11.3	9.3	9.4	14.4	6.0	12.8	13.4	11.9	10.8	8.3	13.0	12.4
SD	4.7	3.5	3.3	3.8	4.2	2.4	3.6	3.4	3.4	3.7	3.2	3.4	4.0	3.4	4.4	4.4

\*All computations based on raw, error scores and ungrouped data.

to go on record as having become increasingly impressed by the complexity of the English language and with its imperfections as a system of communication. For example, he is now able to use a word, with a Thorndike-Lorge *G*-count of 32, with a meaning so far unrecognized by a small sample of his highly verbal colleagues. This means that the Webster synonym dictionary lists frequently used words under meaning categories wherein they are rarely found otherwise. In extension it means that the *G*-count is a highly inaccurate predictor of difficulty in such instances; and that presently unknown levels of difficulty have been built into this instrument.

The language sample assembled by these procedures is believed to be unique. The 80 stem words in the test were sought in Ogden's (3) Basic English vocabulary. Fifty (62.5 per cent) were found in some clear-cut form. For obvious reasons this value cannot be directly compared with similar values from other vocabulary tests. If such comparisons were to be made they should be accomplished on the basis of concept coverage. However, it is suspected that the procedures used essentially insure a more comprehensive concept coverage than would be found in other vocabulary measures.

One further general remark is believed to be appropriate. The various devices used to sort out semantic overlap were designed not only to facilitate the discriminations required, but also to rid the instrument of something which might be called a flexibility-rigidity factor. The writer believes that language behavior variability and flexibility-rigidity are not completely separable, and he is confident that this second variable is to some degree reflected in the Synonym Vocabulary Test score. He has now under way a test construction project designed to maximize the measurement of the flexibility-rigidity factor.

#### N. SUMMARY

The Synonym Vocabulary Test, its rationale, construction, and try-out have been described in detail. The instrument consists of 16 matching-type items. In each item the five stem words were chosen for high frequency of use. The response words were also chosen on the basis of frequency of use and in such fashion as to attempt to maximize the range of difficulty within each item.

In the population used, the test appeared to have a reliability in the neighborhood of .97.

Minor and incidental studies appeared to show that the scores correlate with certain criteria of academic achievement.

Some evidence is presented which appears to show that the test samples language behavior in a somewhat unique fashion.

## REFERENCES

1. HULL, C. L. Principles of behavior. New York: Appleton-Century, 1943.
2. LEPLEY, W. M. Variability as a variable. *J. of Psychol.*, 1954, **37**, 19-25.
3. OGDEN, C. K. The System of Basic English. New York: Harcourt, Brace, 1934.
4. THORNDIKE, E. L., & LORGE, I. The Teachers Word Book of 30,000 Words. New York: Bureau of publications, Teach. Coll., Columbia Univ., 1944.
5. WEBSTER'S DICTIONARY OF SYNONYMS. Springfield, Mass.: Merriam, 1951.
6. WEBSTER'S UNABRIDGED DICTIONARY. Springfield, Mass.: Merriam, 1934.

*Department of Psychology*  
*Pennsylvania State University*  
*State College, Pennsylvania*





## SPATIAL ORIENTATION IN MAN AFTER CEREBRAL INJURY: I. ANALYSES BY LOCUS OF LESION\*<sup>1</sup>

*New York University, Bellevue Medical Center*

JOSEPHINE SEMMES,<sup>2</sup> SIDNEY WEINSTEIN, LILA GHENT, AND  
' HANS-LUKAS TEUBER

### A. INTRODUCTION AND PROBLEM

Numerous case reports attest that disorientation in space may be a prominent sequel of cerebral lesions in man. A great variety of such disturbances has been described, ranging from defective localization of objects in homonymous amblyopic portions of the visual field, to gross misconceptions of place and situation associated with confusional states. Intermediate between these extremes are disorders for which no sufficient basis in sensory or motor impairment, or in general dementia, has been found. They are said to resemble agnosias or apraxias. Among these states are difficulties in estimation of direction and distance of objects (3, 4, 6, 12, 13, 14, 30, 32, 33, 36), neglect of half of the body or of external space (1, 4, 19, 20, 23, 25, 27, 28, 29), confusion of right and left (7, 10, 12, 23, 25, 27, 36), defects in constructional tasks (3, 10, 16, 21, 22, 23, 27, 28, 29, 30, 37), "apraxia" for dressing (4, 23, 29), inability to follow habitual routes (4, 10, 13, 14, 23, 24, 26, 27, 28, 29, 41), and disturbance of place recognition (10, 17, 24, 27, 28, 29, 41).

Although these abnormalities are said to be dissociable from one another, and from object agnosia, the consensus is that lesion of the posterior parietal lobe is responsible for any or all of them. The evidence is equivocal, however, for lesions of the occipital, temporal, and even frontal lobes (9), are sometimes cited as having given rise to such alterations. Discussions of localization in the literature most often center about the rôles of the right and left hemispheres in spatial functions (4, 7, 16, 17, 23, 25, 28, 31, 37), and the relative potency of bilateral *versus* unilateral lesions in the production of syndromes of defective spatial reference (8, 13, 14, 18, 33).

The relation of modality to these disturbances also requires clarification. Most reports have emphasized the disorganization of *visual* space. Poppel-

\*Received in the Editorial Office on August 19, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>The program of research of which this study is a part is supported by the Commonwealth Fund of New York, and by the Office of the Surgeon General (Department of the Army), Contract DA-49-007-MD-312.

<sup>2</sup>Postdoctoral fellow, U. S. Public Health Service, 1952-1954.

reuter described what he called a specifically "optic" apraxia (30) in patients who had recently sustained occipital, temporal, or posterior parietal gunshot wounds. Holmes (13), describing cases in which position of objects was incorrectly perceived, says: "It was only when they relied on sight alone that they could not localize the position of an object in space. If it touched any part of their own bodies, they could always bring their hands to it immediately and correctly; touch gave them the necessary local knowledge that they failed to obtain from vision alone." McFie, Piercy, and Zangwill (23) likewise remark that their Case 4 could tie his tie *except* when regarding himself in a mirror.

Thus it appears that at least some forms of spatial disorder may be restricted to the visual sphere; other forms are also called "visual disorientation" whereas in fact restriction of the defect to visually-guided behavior is far from clear. Not all of Holmes' cases (12) showed clear sensory dissociation in orientational behavior. Case I had difficulty in circumventing obstacles perceived by touch (in contrast to Case VI), and Case V appeared to be deficient in auditory localization. Holmes is cautious in interpreting the nature of the defect in these cases, ascribing it to involvement of "the cerebral mechanisms, by which the spatial percepts received through the different senses are associated and assimilated with past experience."

The majority of subsequent reports by other authors imply, but do not establish, that the defects they describe were modality-specific. A few consider the defect "visual" even though they present evidence that it was more inclusive. Brain (4) states that one of his patients was more disoriented in the dark than in the light, and that he was unable to follow routes or locate objects although visualization was preserved. The defect is nevertheless interpreted as primarily visual. He says: "Severe damage to the central processes of space perception on the visual side deranges the whole basis of orientation, and, as several writers have pointed out, patients with visual disorientation are far more severely handicapped than a blind man, and lack a blind man's capacity to orientate himself by other senses."

Paterson and Zangwill (28) conclude that the disorder in the cases they present was "clearly restricted to visual-spatial activities and functions." In description of their Case I, however, they report that the consistent errors in pointing to objects, which they interpret as one indication of "agnosia" for the left half of space, were made only with the eyes closed, not with the eyes open. They also note that the patient's difficulties in counting objects were not lessened by the use of his forefinger. In some of Holmes' cases (12), by contrast, correct enumeration was possible when touch was allowed.

The interpretations of such constellations of defect as "visual" perhaps overextends the meaning of the term. If the patient, through habit, fails to make use of orientational cues which are available to the blind and depends instead on distorted visual impressions, it is understandable that a truly visual disorientation could be more disabling than blindness, in the same sense that hemiambyopia may be more disabling than hemianopia. However, if defects in orientational behavior persist or become worse when vision is excluded, this evidence appears to disprove the contention that such defects are based on altered visual function. If the contention is not refuted by this evidence, then it is not clear what operational criteria can be used to distinguish between specifically visual and more general defects.

The question of modality-specificity has both theoretical and practical significance. On the theoretical side, further elucidation of the nature of these defects, and ultimately our notions of how the nervous system mediates spatial functions, presuppose knowledge of the degree of independence of modalities in such perceptual organizations. On the practical side, it is important as a guide to rehabilitative procedure to know if the patient can profit by information about the spatial aspects of the environment conveyed by a sensory route other than vision.

The principal source of data on spatial disorientation after cerebral lesions has been case histories. Although the value of careful study of selected individual cases is undeniable, false inferences both as to the nature and localization of symptoms may be drawn from them. Apart from the pitfalls of using clinical material to determine the effects of uncomplicated loss of tissue (11), analyses of the effects of localized dysfunction based upon selected cases are susceptible to several sources of error: (a) The frequency of cases with a given lesion but without given symptoms is not assayed; (b) when lesions are multiple or widespread, the symptoms may be ascribed to some part or element of the total lesion (commonly to the largest of the lesions, thus pressing the principle of mass action into the service of localization); (c) abnormal performance on tests is alleged to exist, without controls or with insufficient report of controls for the reader to judge adequacy of matching; and (d) relative dissociation of symptoms is reported without regard to possible normal inequalities in the functions tested.

Only a few authors (8, 34) have considered the possibility that dissociation of symptoms may appear at different stages of dissolution or restitution of function, so that the constellation of deficits may vary from one examination to another. Such variability is understandable if one considers the nature of the cases studied subsequent to Holmes' and Poppelreuter's re-

ports. Many of these cases suffered from acute neoplasm or vascular accident; the instances of penetrating trauma in which spatial disorientation was a major feature have rarely been followed beyond the early post-traumatic phase.

The least equivocal picture may be expected to result from studies of a large group of patients, tested long after the cerebral lesions were sustained. The minimal but stable deficits which can be found in such cases are likely to reveal whether spatial disorientation is in fact: (a) a predominantly visual disorder, as most authors have assumed; or (b) whether it is a disorder of motor control, specific for skilled acts, i.e., a form of apraxia, as claimed by others (15); or (c) whether it transcends modality, implicating orientation through other senses as well as vision (2, 5).

To decide among these possibilities, we devised an objective test of spatial orientation, and applied it to a large unselected sample of brain-injured cases (from one to 10 years after their injury) and to a control group. The test required the subject to follow routes represented on maps. Some of the maps were perceived visually; other maps, which were rotations and mirror-images of the visual maps, were perceived through touch alone. Thus comparable scores were obtained for visually- and tactually-guided behavior. In the present report, the scores are analyzed by lesion groups, formed according to the major lobes of the brain, laterality, and involvement of one or both hemispheres. A subject became a member of a given lesion group if the evidence indicated that there was *any* damage to the sector in question.

Subsequent reports will disregard locus of lesion and deal exclusively with association or dissociation of symptoms. The brain-injured sample will be subdivided into symptom groups, based on whether or not there was visual or somatosensory defect. Performance on the spatial orientation test will also be evaluated in relation to discriminations of the form and size of objects, and to a variety of measures of orientation with respect to the subject's own body.

## B. MATERIAL AND METHODS

### 1. Subjects

The subjects were drawn from the population of male veterans available to the Psychophysiological Laboratory of New York University College of Medicine. These men have sustained penetrating injuries of the brain or of peripheral nerves in World War II or the Korean campaign. The majority of them have participated in previous studies from this laboratory (35, 38, 39, 40). All those used in the present study are ambulatory, and came voluntarily for individual testing sessions. The experimental group



comprised 62 men with loss of cerebral tissue, established by review of surgical and roentgenographic records. No bias directed our choice of subjects from the total available population except that we excluded men whose sensory or motor impairment was sufficiently severe, in their judgment or ours, to raise doubts as to the adequacy of their sensory experience with the tactual maps. The control group was composed of 17 men: 12 had injuries to one or more of the peripheral nerves supplying the leg; the remaining five had injuries to the radial or ulnar nerves. The injuries to the control group were assumed to be without significance for the ability tested.

The distribution of the experimental group by locus of lesion is shown in Table 1. The evidence regarding mass and kind of tissue destroyed was not sufficiently clear to warrant distinctions on these bases; in all cases, at least cortex was damaged, and we have been able to estimate only the minimal extent of destruction. The method of forming the lesion groups is outlined in detail in a recent publication (35). A subject was assigned to the lesion groups corresponding to his injury, as represented on standard views of the cerebral hemispheres. Each grouping was made independently of the others. For example, an individual with damage to the right parietal and left occipital lobes became a member of the Parietal group, the Occipital group, the Right group, the Left group, and the Bilateral group. In order to avoid comparisons of groups which were not mutually exclusive, the performance of each lesion group was compared to that of a complementary group consisting of all brain-injured subjects *without* damage to the sector under consideration. Therefore, the subject mentioned above also belonged to the Non-Frontal group and to the Non-Temporal group.

Seven separate analyses of variance of the resulting scores were made, in outline as follows:

1. Controls *versus* Frontals *versus* Non-Frontals.
2. Controls *versus* Parietals *versus* Non-Parietals.
3. Controls *versus* Temporals *versus* Non-Temporals.
4. Controls *versus* Occipitals *versus* Non-Occipitals.
5. Controls *versus* Bilaterals *versus* Unilaterals.
6. Controls *versus* Lefts *versus* Non-Lefts.
7. Controls *versus* Rights *versus* Non-Rights.

In each of these analyses, appropriate error-terms were derived by dividing the variance within subclasses into two parts, one part ascribable to the total scores (the sums of the scores on the visual and the two tactual sub-tests) of individuals within groups (76 degrees of freedom), and a remainder (152 degrees of freedom). Group differences were evaluated by means of the

TABLE 1A  
DISTRIBUTION OF LESIONS BY LOBE AND LATERALITY IN 62 BRAIN-INJURED MEN

Unilateral lesions				Bilateral lesions			
No.	L	No.	R	No.	L	R	
6	F	5	F	9	F	F	F
4	P	2	P	1	F	FP	FP
2	T	3	T	1	F	FT	FT
3	O	3	O	1	F	T	T
				1	FP	F	F
1	FP	1	FP	1	FP	FP	FP
1	FT	3	FT	1	FPTO	F	F
				1	PO	PO	PO
2	PO	3	PO	1	PTO	TO	TO
1	TO	1	TO	1	PTO	O	O
1	PTO	1	PT	1	O	PTO	PTO
Totals: 21		22		19			

F = Frontal. P = Parietal. T = Temporal. O = Occipital. R = Right Hemisphere. L = Left Hemisphere. No = Number of Cases.

TABLE 1B  
SUMMARY GROUPING OF 62 BRAIN-INJURED CASES BY LOCUS OF LESION

Unilateral lesions				Bilateral lesions			
Lobe	Restricted to indicated lobe	Plus other lobes in same hemisphere	Total	Bilateral for indicated lobe	Unilateral for indicated lobe plus other lobes in opposite hemisphere	Total	TOTAL
Frontal	11	6	17	14	1	15	32
(Non F)	—	—	26	—	—	4	30
Parietal	6	9	15	2	6	8	23
(Non P)	—	—	28	—	—	11	39
Temporal	5	8	13	1	5	6	19
(Non T)	—	—	30	—	—	13	43
Occipital	6	8	14	4	1	5	19
(Non O)	—	—	29	—	—	14	43

former, whereas the latter was used to assess the significance of differences among modes of perceiving the maps, and of the interaction of group and mode.

## 2. Description of the Test

The test employed 15 diagrams of paths, as shown in Figure 1. The nine dots on each of the maps represented nine circular red spots, 15.3 cm. in diameter, painted on the gray floor of a large room. Adjacent spots were

137 cm. apart, center to center. The subject was given each map in turn, in the order shown in Figure 1, and required to walk through the indicated path. The direction "North" was shown on the maps (at the top in Figure

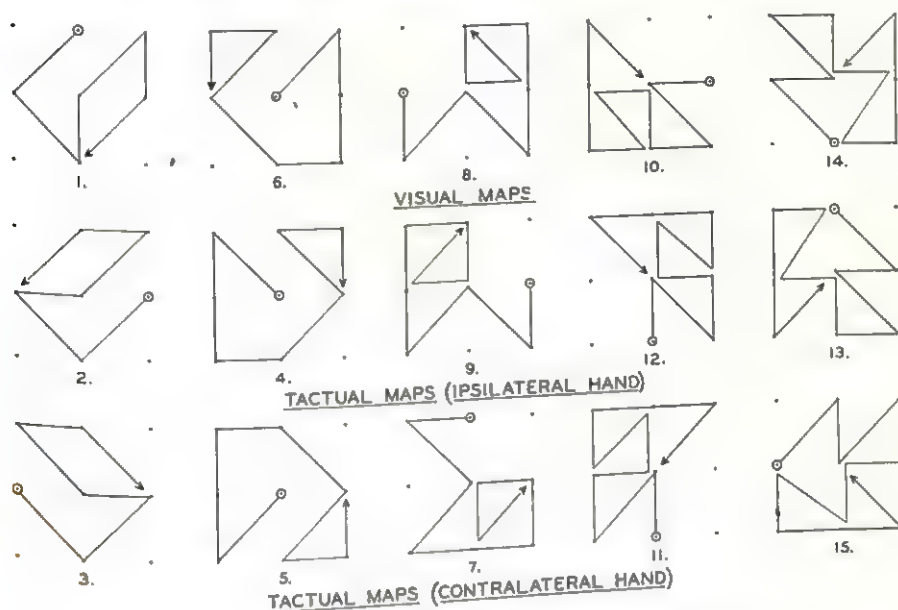


FIGURE 1

1), and the subject was informed which was the north wall of the room. He was instructed to keep the map before him as he progressed along the route, and to hold it in a constant orientation with respect to his body, that is, with the southern edge of the map always nearest him. Thus the map, more often than not, was incorrectly oriented to actual directions as the subject made the indicated turns; he therefore had to effect a translation of the spatial coordinates of the map into the corresponding real coordinates in order to reproduce the path correctly.

All the maps were square plaques, 30.5 cm. on a side. Five of the maps were made of white cardboard on which the path was drawn in black lines 4.5 mm. in width. The nine dots (each 12 mm. in diameter) were made with red paint. A black circle around one of the dots indicated the starting point, and an arrowhead marked the end. These five maps were presented for visual inspection.

The remaining 10 maps were perceived tactually, with vision excluded. The subject reached underneath a black curtain into a shallow cardboard box, which held the map. He carried the box with one hand and felt the map

with the other. The tactual maps were made of plywood, with the nine spots represented by the heads of upholstery tacks, 10 mm. in diameter. The path was indicated by a continuous cord, 4.8 mm. in diameter, stapled to the plywood at the turning points. A knot around one of the tacks marked the starting point of the route, and a cardboard arrowhead, the end.

In the experimental group, five of the tactual maps were presented to the hand contralateral to the brain lesion (or to the larger lesion, in the bilateral cases), and five to the ipsilateral hand. In the control subjects with radial or ulnar nerve injury, the affected hand received the maps designated for the "contralateral" hand, and the unaffected hand, those designated for the "ipsilateral" hand. Of the control cases with leg injuries, half used the right hand and half the left hand as the "contralateral" hand.

The instructions to the subject, as he was given the first map (a visual one), were approximately as follows:

This is a test which we hope will tell us how good your sense of direction is. Notice these nine red spots on the floor. Here is the first of a series of maps on which these points are represented. You will be expected to walk along a path on the floor corresponding to that marked out on the map. Start from the circled point, and walk from one point to the next and so on until the path is completed, here at the arrowhead. "North" is marked on the map, and this way (pointing) is north in this room. Hold the map in front of you, with its northernmost edge farthest away, and keep it in that relation to you, even though it won't always be in the proper orientation to the room. Walk directly forward—do not walk backward or sidewise. When you reach a turning point, do not turn until you have decided to which point you are going next.

The second map was a tactual one, and its makeup was described to the subject as it was given to him. The general instructions remained the same as for the visual maps. Any additional explanations requested by the subject were given, with regard to either the visual or tactual maps.

The path followed by the subject in response to each of the maps was called "right" or "wrong," as a whole. False turnings were not considered errors unless deviation from the correct sequence of points occurred, that is, the subject was allowed to correct an incorrect turn if he did so before reaching the next point. The maps were given weighted scores, based on the number of turns required. Thus the first map in each of the three series (visual, tactual ipsilateral, and tactual contralateral) was scored five; the second map six; and so on to the last three maps, one from each series, which were scored nine each. The subject received a total score for each series,



which was the sum of the points earned by his performances on the five constituent maps.

### C. RESULTS

The principal finding was that the performance of the group with lesion of the parietal lobe was significantly inferior to that of the control group (.01 level of confidence) and to that of the brain-injured group without parietal lesion (.05 level of confidence). The mean score of the group of subjects with lesions in parts of the brain other than the parietal lobe was not significantly different from the control group mean. No other division of the total brain-injured sample revealed significant differences among groups. Table 2 summarizes the analyses of variance. Figures 2, 3, 4, 5, and 6 show the mean scores by groups, by modes of perceiving the maps, and by groups within each mode.

TABLE 2  
SUMMARY OF ANALYSES OF VARIANCE

Each column shows the values of *F* derived from a separate analysis: 1 = Controls vs. Frontals vs. Non-Frontals; 2 = Controls vs. Parietals vs. Non-Parietals; 3 = Controls vs. Temporals vs. Non-Temporals; 4 = Controls vs. Occipitals vs. Non-Occipitals; 5 = Controls vs. Bilaterals vs. Unilaterals; 6 = Controls vs. Lefts vs. Non-Lefts; 7 = Controls vs. Rights vs. Non-Rights.

Source	1	2	3	4	5	6	7
Groups	2.80	4.15*	1.81	1.89	1.89	1.79	1.95
Modes	2.38	2.41	2.37	2.37	2.41	2.42	2.43
Interaction	.68	1.27	.56	.59	1.14	1.38	1.49

\*Significant between the .05 and .01 levels of confidence.

The experimental design adopted (use of the same subjects in all three modes) rendered the analyses maximally sensitive to differences in score attributable to modes of perceiving the maps, and to group differences in response to these modes. Nevertheless, these differences did not reach significant levels. Neither mode nor the interaction of group and mode proved to be a significant source of variation in performance. Likewise, the correlations among scores made on the three series of maps were significantly positive in all groups (see Table 3).

### D. DISCUSSION

The results indicate that the objective test of spatial orientation employed in this study measures an ability which can be affected by brain injury. This effect persists well beyond the acute post-injury phase. Moreover, performance is not affected equally by injuries in different cerebral loci. The

MEAN SCORES BY GROUPS, ALL MODES COMBINED

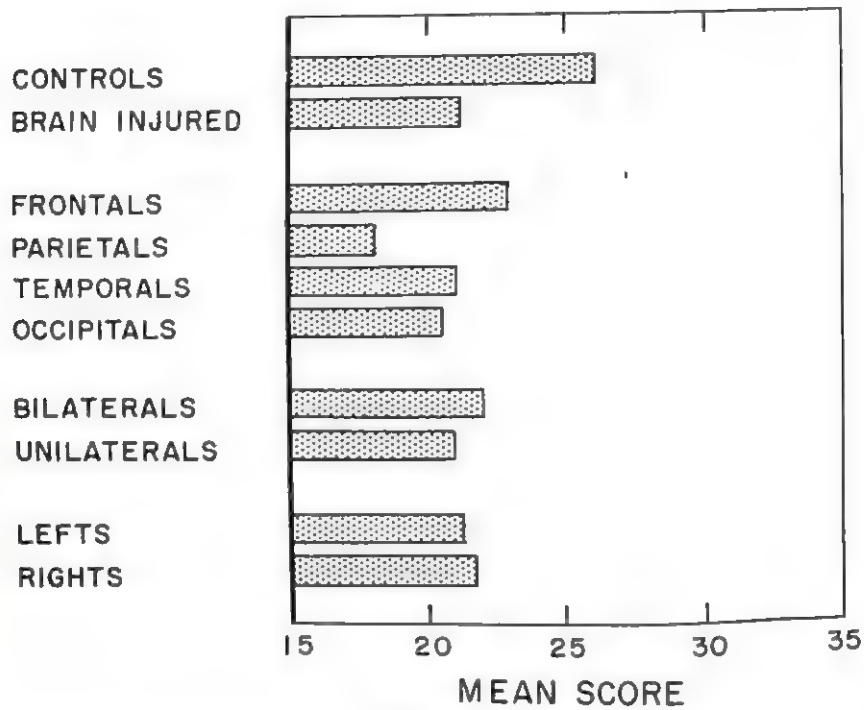


FIGURE 2

MEAN SCORES BY MODES, ALL GROUPS COMBINED

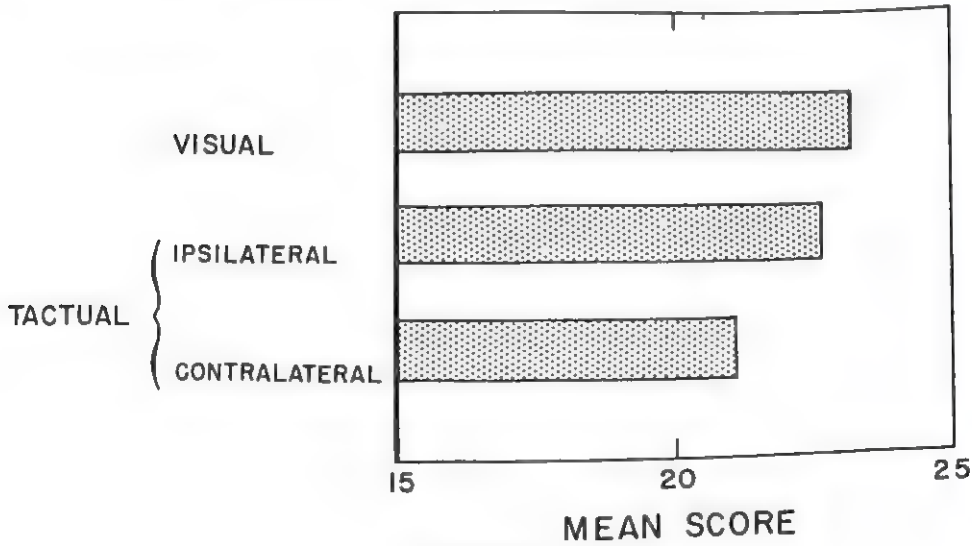


FIGURE 3

TABLE 3  
CORRELATIONS BETWEEN SCORES DERIVED FROM THE MAPS PERCEIVED VISUALLY (V),  
TACTUALLY WITH THE HAND IPSILATERAL TO THE LESION ( $T_1$ ), AND  
TACTUALLY WITH THE HAND CONTRALATERAL TO THE LESION ( $T_c$ )

	$r_{VT_1}$	$r_{VT_c}$	$r_{T_1T_c}$
Controls	.521*	.482*	.731**
Frontals	.586**	.451**	.579**
Non-Frontals	.742**	.746**	.811**
Parietals	.687**	.587**	.805**
Non-Parietals	.650**	.588**	.586**
Temporals	.758**	.704**	.798**
Non-Temporals	.640**	.564**	.679**
Occipitals	.774**	.660**	.830**
Non-Occipitals	.628**	.572**	.662**
Bilaterals*	.655**	.473*	.717**
Unilaterals	.674**	.670**	.704**
Lefts	.610**	.610**	.728**
Non-Lefts	.816**	.580**	.720**
Rights	.659**	.595**	.698**
Non-Rights	.723**	.614**	.730**

The hands were designated "ipsilateral" and "contralateral" with respect to the larger lesion in the cases with bilateral injury.

\*Significant at or beyond the .05 level of confidence.

\*\*Significant at or beyond the .01 level of confidence.

integrity of the parietal lobe is most important to this ability, whereas preservation of other regions appears to be less important. These results are in harmony with the classical localization of the syndrome of spatial disorientation.

Comparison between the results of this study and of previous ones is necessarily tentative and approximate. We have used a different measure of spatial orientation from those used in clinical practice, and have studied a brain-injured population which differs from those generally used in type of lesion and in time since injury. Our conclusions may therefore have limited relevance to the forms of disability discussed in the literature. In support of the validity of our test, however, we might adduce the following evidence: Impairment in route-following, and in the ability to make maps and ground plans, are among the most commonly cited elements of the clinical syndrome of spatial disorientation; men who complained that the injury had affected their "sense of direction" were among those who made the poorest scores on the test; as mentioned previously, the localization of the lesion producing maximal impairment on the test conforms, at least in terms of lobe injured, to the general view of locus of the causative lesion in spatial disorientation.

Perhaps the most important difference between our results and those reported in the literature lies in the nature of the defect. Disorientation in space has generally been considered to affect only the organization of visual perception, or the expression of visual perception in appropriate action. If

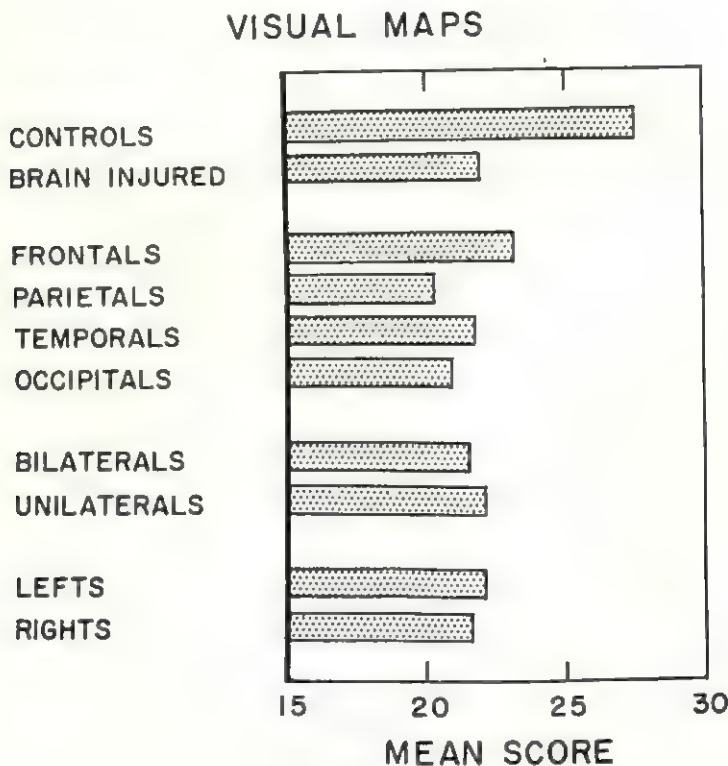


FIGURE 4

either of these conceptions properly describes how spatial ability is affected by parietal injury, our group with lesion of this region should have shown depression of scores on the visual maps only. Analyses of the data, however, do not allow us to conclude that any mode was more affected than the others, and in fact the Parietal group showed a slightly greater percentage decrement on the tactual maps than on the visual series.

A somewhat different view of the manner in which brain injury affects orientation in space has also been proposed (8, 25). Disorientation, resulting from a lesion in the dominant parieto-occipital region, is said to manifest itself as principally visual, principally somesthetic, or principally in the sphere of language, depending on the proximity of the lesion to the occipital pole,



the parietal region, or Wernicke's zone, respectively. According to this conception, we should expect to find, among our cases of parietal and occipital injury, some who showed isolated decrement in scores on the visual maps and others who showed deficit confined to the tactual maps. Perhaps it

### TACTUAL MAPS (IPSILATERAL HAND)

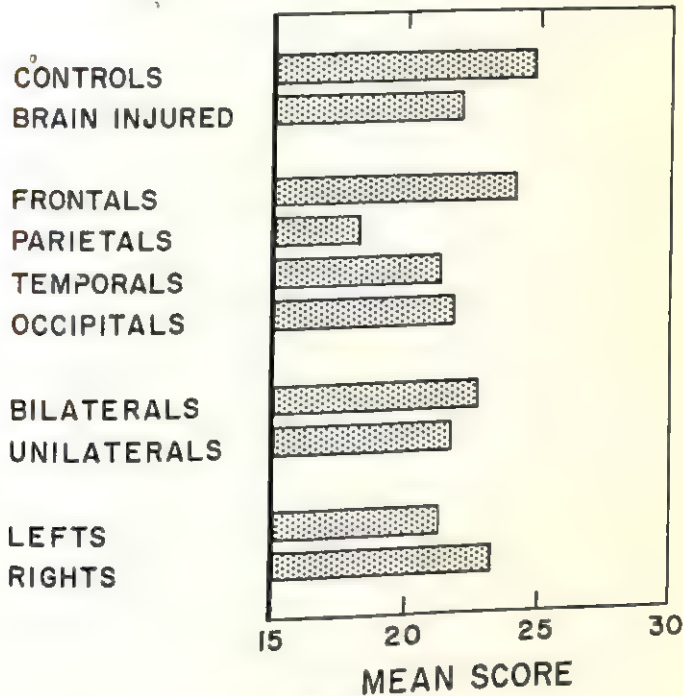


FIGURE 5

might be urged that the use of grouped data cancels out and obscures the effects of such dissociations. However, the substantial correlations we found between scores on the visual and the tactual maps in the Parietal and Occipital groups, as well as in all other groups, lend no support to this view. We have also examined the data to discover whether there were individuals whose scores suggested impairment limited to one modality. Specifically, we looked for scores in the upper third of the range on one mode associated in the same individual with scores in the lower third on another mode. Out of the 62 brain-injured subjects we tested, there was only one who showed isolated depression of scores in the visual maps. This subject had a bilateral lesion of the frontal lobes. There were two subjects who made low scores on one

or both tactual series together with high scores on the visual series. Both these subjects had unilateral lesions overlapping the central fissure, one in the right hemisphere extending backward into the posterior parietal region, and the other in the left hemisphere extending forward into the prefrontal zone. Neither of these cases had a history of left-handedness.

### TACTUAL MAPS (CONTRALATERAL HAND)

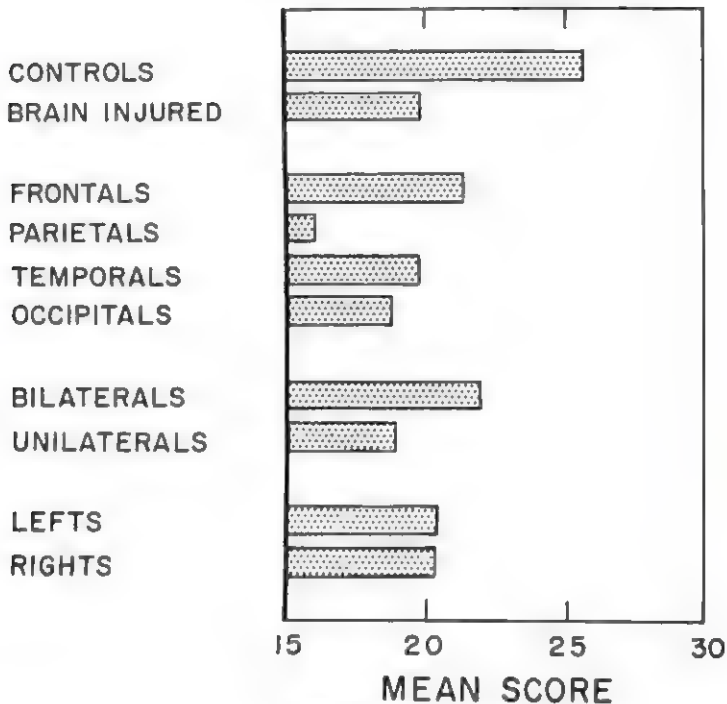


FIGURE 6

The objection might be raised that since the response we measured in all three series was locomotion in a space perceived visually (i.e., the spots on the floor of the room), this visual element common to all three series could account for poor performance on the tactual maps as well. This view implies that the perceived spatial relationship of the spots on the floor was faulty, and consequently hindered execution of paths which were perceived accurately by touch. The character of the subjects' performance, and in particular their errors in performance, makes this interpretation improbable. None of the subjects had any difficulty in walking directly to an indicated point, and none showed any past-pointing or other dysmetric phenomena in

locomotion. Most of the errors made involved confusions of the path by  $180^\circ$  at a choice point, or failures to follow the path represented on the map continuously, i.e., substitution of some earlier or later section of the path which touched the same point. These considerations likewise suggest that the impairment was not an "apraxia," by any of the classical definitions of the term, for the errors were acts which were entirely comparable to correct responses in coordination and complexity. In none of the subjects' comments on performance was there a suggestion of a discrepancy between intention and locomotor behavior.

The possible dependence of our results on receptive defects must be considered. Another report will deal with these relationships in detail, but it is appropriate to examine them briefly here. It is clear that defects confined to one modality cannot explain an impairment which cuts across modalities. However, it is conceivable that a chance coincidence, in our material, or visual and somatosensory defects in the same individuals might underlie the correspondence of scores on the visual and tactual maps. Fifteen of the brain-injured subjects had visual field defects; 31 had defects in somatosensory function of the hand (appreciation of the direction of passive movement, tactual sensitivity, two-point discrimination, or point localization). Only six individuals were common to these two groups. Five of these cases were included in the Occipital group (25 per cent), which was not significantly different from the control group in score; four of them (17 per cent) were included in the Parietal group. Such a small number of cases, so distributed, could not have determined the outcome. \*

From the evidence here presented, it appears unlikely that disorders of spatial perception which persist beyond the early post-traumatic period are dissociable along modality lines. It may be significant that the two of Holmes' cases who were observed after the first three months following injury (Cases I and V) were also those who showed orientational defects which were not restricted to visually-guided behavior. The forms of disorientation which are manifest only in the visual sphere may be transient phenomena seen primarily in the acute phase after cerebral insult or as a concomitant of irritative processes. Bay's suggestion (2) is plausible that such forms are secondary to changes in primary visual function, undetected by perimetry and other standard methods of examination. Such disorders may bear no relation to the chronic disturbance.

The chronic disturbance affects spatial patterns conveyed by touch as well as by vision. It is therefore not properly described as an agnosia. We have also presented the reasons for our belief that it is not apraxic in nature.

The class of defect to which it belongs may be suggested by analyses of its relation to other abilities, and in particular, to other measures of spatial perception.

#### E. SUMMARY

1. An objective test of spatial orientation which results in comparable scores for visually- and tactually-guided behavior is described. The test required the subject to follow by locomotion routes represented on maps. Five of the path-diagrams were perceived visually; the remaining 10, five for each hand, were perceived by touch alone. The tactual maps were rotations and mirror images of the visual ones.

2. The test was given to 62 subjects with traumatic injury to the brain, and to 17 control subjects. The results were analyzed according to the cerebral lobe injured, laterality of lesion, and involvement of one or both hemispheres.

3. The mean score of the group with lesion of the parietal lobe was significantly inferior to that of the control group, and to that of the complementary brain-injured group with lesions which did not encroach upon the parietal lobe. The latter group showed no significant decrement in performance relative to the controls.

4. Differences in performance ascribable to the mode of perception of the maps (visual, tactual via the hand ipsilateral to the lesion, or tactual via the hand contralateral to the lesion) were insignificant. Likewise, the interaction of group and mode did not contribute significantly to the variation of performance. Correlations among the scores on the three series of maps were significantly positive in all groups.

5. These results do not substantiate the frequent claim in the literature that disorders of route-following reflect disturbance of a high level of specifically visual function. On the contrary, whether the modality to which the stimuli are presented is vision or touch appears to be of little importance. Since the disorder is neither specific for modality nor for executive function, it should not be described as an agnosia or apraxia.

#### REFERENCES

1. ALAJOUANINE, T., THUREL, R., & OMBREDANE, A. Somato-agnosie et apraxie du membre superieur gauche. *Rév. Neurol.*, 1934, 1, 695-703.
2. BAY, E. Disturbances of visual perception and their examination. *Brain*, 1953, 76, 515-550.
3. BENDER, M. B., & TEUBER, H. L. Spatial organization of visual perception following injury to the brain. *Arch. Neurol. & Psychiat.*, 1948, 59, 39-62.
4. BRAIN, W. R. Visual disorientation with special reference to lesions of the right cerebral hemisphere. *Brain*, 1941, 64, 244-272.



5. CONRAD, K. Versuch einer psychologischen Analyse des Parietalsyndroms. *Monatschr. Psychiat. Neurol.*, 1932, 84, 28-97.
6. GELB, A. Die psychologische Bedeutung pathologischer Störungen der Raumwahrnehmungen. *Kongr. Exper. Psychol.*, 1925. Jena: Fischer, 1926.
7. GERSTMANN, J. Fingeragnosie: Eine umschriebene Störung der Orientierung am eigenen Körper. *Wien. klin. Wschr.*, 1924, 37, 1010-1012.
8. GOLDSTEIN, K. Die Lokalisation in der Grosshirnrinde. *Hdb. norm. pathol. Physiol.* (Ed. A. Bethe et al.), 1927, 10, 600-842. (See esp. 805-813.)
9. HARTMANN, F. *Die Orientierung*. Leipzig: Vogel, 1902.
10. HEAD, H. *Aphasia and Kindred Disorders of Speech*. Cambridge: Cambridge Univ. Press, 1926.
11. HEBB, D. O. Man's frontal lobes. *Arch. Neurol. & Psychiat.*, 1945, 54, 10-24.
12. HOLMES, G. Disturbances of visual orientation. *Brit. J. Ophthalm.*, 1918, 2, 449-468, 506-516.
13. ———. Disturbances of visual space perception. *Brit. Med. J.*, 1919, 2, 230-233.
14. HOLMES, G., & HORRAX, G. Disturbances of spatial orientation and visual attention, with loss of stereoscopic vision. *Arch. Neurol. & Psychiat.*, 1919, 1, 385-407.
15. KLEIN, R. Über die Funktionen des Parietallappens. *Nervenarzt*, 1933, 6, 1-7, 67-74.
16. KLEIST, K. *Gehirnpathologie*. Leipzig: Barth, 1934.
17. LANGE, J. Agnosien und Apraxien. *Hdb. Neurol.* (Ed. O. Bumke and O. Foerster), 1936, 6, 807-960. Berlin: Springer, 1936.
18. LHERMITTE, J., & TRELLES, J. O. Sur l'apraxie pure constructive. Les troubles de la pensée spatiale et de la somatognosie dans l'apraxie. *Encéphale*, 1933, 28, 413-444.
19. IVES, E. R., & NIELSEN, J. M. Disturbance of body scheme: Delusion of absence of part of body in two cases with autopsy verification of lesion. *Bull. Los Angeles Neurol. Soc.*, 1937, 2, 120-125.
20. MARTIN, J. P. Consciousness and its disturbances. *Lancet*, 1949, 256, 48-53.
21. MAYER-GROSS, W. Some observations on apraxia. *Proc. Roy. Soc. Med.* (Sect. Neurol.), 1935, 28, 63-72.
22. ———. The question of visual impairment in constructional apraxia. *Proc. Roy. Soc. Med.* (Sect. Psychiat.), 1936, 29, 66-70.
23. MCFIE, J., PIERCY, M. F., & ZANGWILL, O. L. Visual-spatial agnosia associated with lesions of the right cerebral hemisphere. *Brain*, 1950, 73, 167-190.
24. MEYER, O. Ein- und doppelseitige homonyme Hemianopsie mit Orientierungsstörungen. *Monatschr. Psychiat. Neurol.*, 1900, 8, 440-456.
25. NIELSEN, J. M. Gerstmann syndrome: Finger agnosia, agraphia, confusion of right and left and acalculia. Comparison of this syndrome with disturbance of body scheme resulting from lesions of the right side of the brain. *Arch. Neurol. & Psychiat.*, 1938, 39, 536-560.
26. NIELSEN, J. M., & OLSEN, C. W. Mind-blindness. Optical disorientation in space and simultanagnosia. *Bull. Los Angeles Neurol. Soc.*, 1936, 1, 73-81.
27. NIELSEN, J. M., & SULT, C. W., JR. Agnosias and the body scheme. *Bull. Los Angeles Neurol. Soc.*, 1939, 4, 69-76.
28. PATERSON, A., & ZANGWILL, O. L. Disorders of visual space perception associated with lesions of the right cerebral hemisphere. *Brain*, 1944, 67, 331-358.
29. ———. A case of topographical disorientation associated with a unilateral cerebral lesion. *Brain*, 1945, 68, 188-212.

30. POPPELREUTER, W. Die psychischen Schädigungen durch Kopfschuss im Kriege 1914-1916: I. Die Störungen der niederen und höheren Schleistungen durch Verletzungen des Okzipitalhirns. Leipzig: Voss, 1917. (See pp. 218 ff., Die optische Apraxie).
31. PÖTZL, O. Zum Apraxieproblem. *Jb. Psychiat. Neurol.*, 1937, **54**, 133-149.
32. RIDDOCH, G. Dissociations of visual perceptions due to occipital injuries, with especial reference to appreciation of movement. *Brain*, 1917, **40**, 15-37.
33. ———. Visual disorientation in homonymous half-fields. *Brain*, 1935, **58**, 376-382.
34. SCHLESINGER, B. A study on dissociation and reorganization of cerebral function. *Confinia Neurol.*, 1941, **4**, 14-31.
35. SEMMES, J., WEINSTEIN, S., GHENT, L., & TEUBER, H. L. Performance on complex tactual tasks after brain injury in man: Analyses by locus of lesion. *Amer. J. Psychol.*, 1954, **67**, 220-240.
36. SMITH, S., & HOLMES, G. A case of bilateral motor apraxia with disturbance of visual orientation. *Brit. Med. J.*, 1916, **1**, 437-441.
37. STRAUSS, H. Über konstruktive Apraxie. *Monatschr. Psychiat. Neurol.*, 1924, **56**, 65-124.
38. TEUBER, H. L., BATTERSBY, W. S., & BENDER, M. B. Performance of complex visual tasks after cerebral lesions. *J. Nerv. & Ment. Dis.*, 1951, **114**, 413-429.
39. TEUBER, H. L., & WEINSTEIN, S. Performance on a formboard-task after penetrating brain injury. *J. of Psychol.*, 1954, in press.
40. WEINSTEIN, S. Weight judgment in somesthesia after penetrating injury to the brain. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 31-35.
41. ZANGWILL, O. L. Discussion on parietal lobe syndromes. *Proc. Roy. Soc. Med.* (Sect. Neurol.), 1951, **44**, 343-346.

*Department of Psychiatry and Neurology*  
*College of Medicine, New York University*  
*477 First Avenue*  
*New York City 16*

\$20.00 per annum  
\$10.00 per volume  
\$5.00 per half volume

QUARTERLY  
Two volumes per year  
Immediate publication

April, 1955  
Volume 39, Second Half

(Founded in 1935 by Carl Murchison)

# The Journal of PSYCHOLOGY

The General Field of Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a bilateral contract, and may not be resold for a period of five years from the date of publication.

APRIL, 1955

(Manuscripts are printed in the order of final acceptance)

An alternative approach in cross-cultural intelligence testing . . . . .	247
By E. TERRY PROTHRO	
Intra-communication and attitude: A methodological note . . . . .	253
By W. GODFREY COBLINER	
Personalities in faces: IV. A descriptive analysis of the perception of women's faces and the identification of some physiognomic determinants . . . . .	269
By PAUL F. SECORD AND JOHN E. MUTHARD	
The space-visualization factors related to temperament traits . . . . .	279
By ERNEST S. BARRATT	
Gradual strengthening of S-R connections or increasing number of S-R connections . . . . .	289
By VIRGINIA W. VOEKS	
Mass action effects in learning . . . . .	301
By A. ROBERT ROLLIN	
Some "first" sources of sex information as reported by sixty-seven college women . . . . .	321
By HENRY ANGELINO AND EDMUND V. MECH	

(OVER)

Copyright, 1955, by The Journal Press  
Provincetown, Massachusetts, U. S. A.

Entered as second-class matter January 26, 1937, at the post-office at Provincetown, Mass., under the Act of March 3, 1879

The authoritarian personality as a stereotype . . . . .	325
BY BERNARD A. STOTSKY AND SHELDON J. LACHMAN	
The effects of praise and reproof on the generalization of learned concepts . . . . .	329
BY TATANIA JUZAK	
Stimulus-discriminability and concept-attainment: A question arising from Baum's experiment . . . . .	341
BY EDNA HEIDBREDER	
Drawings of a three-dimensional object by mental patients: A preliminary report . . . . .	351
BY C. W. CRANNELL AND EDIKA PLAUT	
Transfer of skill on a following tracking task as a function of task difficulty (target size) . . . . .	355
BY RUSSEL F. GREEN	
Intensive equivalences for sucrose and NaCl solutions . . . . .	371
BY J. G. BEEBE-CENTER, M. S. ROGERS, AND W. H. ATKINSON	
Lysergic acid diethylamide (LSD-25): IV. Effect on attention and concentration . . . . .	373
BY M. E. JARVIK, H. A. ABRAMSON, AND M. W. HIRSCH	
Sex stereotypes and acceptance of sex rôle . . . . .	385
BY SOLOMON DIAMOND	
Chemical influences on behavior: I. The effects of a small dose of hyoscine on performance . . . . .	389
BY S. J. ALEXANDER, M. COTZIN, AND G. R. WENDT	
Studies of motion sickness: X. Experimental proof that aviation cadets tell the truth on motion sickness history questionnaires . . . . .	403
BY S. J. ALEXANDER, M. COTZIN, C. J. HILL, JR., E. A. RICCIUTI, AND G. R. WENDT	
Studies of motion sickness: XIII. The effects of sickness upon rifle target shooting . . . . .	411
BY S. J. ALEXANDER, M. COTZIN, J. B. KLEE, AND G. R. WENDT	
Studies of motion sickness: XIV. Subjective reports of the apparent path of motion on a vertical accelerator . . . . .	417
BY M. COTZIN, C. J. HILL, JR., AND G. R. WENDT	
Studies of motion sickness: XVII. The effects of temperature, posture, and wave frequency upon sickness rates . . . . .	423
BY CARSON JOHNSON AND G. R. WENDT	
Lysergic acid diethylamide (LSD-25): V. Effect on spatial relations abilities . . . . .	435
BY H. A. ABRAMSON, M. E. JARVIK, M. W. HIRSCH, AND A. T. EWALD	
Lysergic acid diethylamide (LSD-25): VI. Effect upon recall and recognition of various stimuli . . . . .	443
BY M. E. JARVIK, H. A. ABRAMSON, AND M. W. HIRSCH	
Lysergic acid diethylamide (LSD-25): VII. Effect upon two measures of motor performance . . . . .	455
BY H. A. ABRAMSON, M. E. JARVIK, AND M. W. HIRSCH	
Lysergic acid diethylamide (LSD-25): VIII. Effect on arithmetic test performance . . . . .	465
BY M. E. JARVIK, H. A. ABRAMSON, M. W. HIRSCH, AND A. T. EWALD	
An apparatus for research in human selective learning . . . . .	475
BY CLYDE E. NOBLE AND FRANCIS J. FARESE	
1923-1953 and 20-60 age changes in moral codes, anxieties, and interests, as shown by the "X-O Tests" . . . . .	485
BY S. L. PRESSEY AND A. W. JONES	



## AN ALTERNATIVE APPROACH IN CROSS-CULTURAL INTELLIGENCE TESTING\*

*Department of Psychology, American University of Beirut*

E. TERRY PROTHRO<sup>1</sup>

### A. INTRODUCTION

There is abundant evidence that scores on intelligence tests are influenced by cultural and class factors. Inasmuch as only three or four per cent of the world's population is made up of middle-class Americans, some psychologists have attempted to construct tests of somewhat broader applicability than those in common use in the United States. These tests are called "culture free," and their creators attempt to eliminate from them items which are more readily answered by individuals of one culture than by those of another culture. Although our knowledge of cognitive differences among peoples is woefully inadequate, it is sufficient to cause one to suspect that thoroughgoing adherence to such a criterion of item exclusion would yield an itemless test. It is difficult to conceive of an intelligence test item which would not involve factors of speed, spatial ability, perseveration, competitive enthusiasm, numerical ability, attention to detail or some other ability which some cultures emphasize or neglect. As Anastasi and Foley have pointed out (1, p. 258), intellectual skills are influenced by experience and it is useless to blame the tests for reflecting this fact, Turnbull (4) seems to hold a similar view.

An intelligence test can tell us an individual's present stage of development with respect to certain abilities, and if we are concerned primarily with those abilities, then tests of the usual type should be useful as predictors. If our criterion of success for persons of all cultures is adequate performance in an American academic situation, then intelligence tests of the American type should be effective in other cultures, for these tests are loaded with factors relating to such scholastic achievement. It would follow from this reasoning that successful cross-cultural testing requires that we select our criterion of "intelligent behavior" and construct tests which predict that behavior. As Lucas (2) has indicated, the alternative to a culture-free test

\*Received in the Editorial Office on October 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.  
<sup>1</sup>This study was made possible by a grant from the Rockefeller Brothers Fund to the American University of Beirut. We wish to express our appreciation to the Fund.

is one which focuses on prediction of criteria. Spiker and McCandless (3) have also argued for emphasis on predictive efficiency. Such an approach would not demand equal mean scores of groups with different experiences, and it would not reveal any information about inherent superiorities of the groups studied. It would, however, yield useful information about the possible occurrence of the "intelligent behavior," and of any other related behavior, in the immediate future. If a single test is to be useful in several different cultures, then the criterion must be common to those cultures. Such a cross-cultural test would be "criterion-bound" rather than "culture-free."

### B. THE PROBLEM

The setting of the American University of Beirut offers unusual opportunity for cross-cultural study of testing. The students are largely Arab, and most of them come from Lebanon, Syria, Jordan, and Iraq. Arabic is the native tongue of the area, but students must also have proficiency in English before being admitted to the university. The university is chartered under the laws of the State of New York. The administration is American and the faculty is made up largely of English-speaking Arabs, Americans, and Englishmen. The organization of the university is similar to that of American universities in the United States. There is a College of Liberal Arts, School of Agriculture, a School of Engineering, and a Medical School. With respect to textbooks, course organization, examinations, etc., the school is typically American. Many of the graduates of the university do graduate work in the United States.

In attempting to predict success at this university, then, we would be using a criterion similar to the criteria used in many American studies. The crucial question is whether an American group test of verbal intelligence would be effective with students drawn from Arab culture. A related question is whether a "verbal" test in one language will predict results in academic work conducted in another language.

### C. THE TESTS

A group intelligence test modeled after the Henman-Nelson Test of Mental Ability was constructed in Arabic.<sup>2</sup> The Arabic test was so similar to the Henman-Nelson Test that it might at first glance seem to be a translation of that test. The actual items, however, were not always mere translations. Vocabulary questions were modified on a basis of preliminary item

---

<sup>2</sup>Construction of the test was done by Miss Balkis Awad under the supervision of Dr. Habib Kurani.

analysis, and some of the information items were changed. It was also found desirable to extend the time limit of the test to one hour so that the students would answer enough items to give the scores reliability. The final forms of the test consisted of 100 multiple choice items involving questions of numerical relations, general information, vocabulary, and spatial relations. This test was called an Arabic Test of Mental Ability.

A 100-item multiple choice vocabulary test was constructed by selecting as stimulus word the first suitable word on every page of an Arabic dictionary, and devising four alternatives for each word, of which one alternative was a synonym for the stimulus word. Words which were more familiar in some dialects than others were not used. Classical words no longer in wide use were eliminated. The test was revised on a basis of preliminary testing and item analysis. In the final form of the test, the items were arranged in approximate order of difficulty.<sup>3</sup>

#### D. PROCEDURE

During the first month of the Fall term, 1952, members of the Freshman class were assembled and administered a battery of aptitude tests. One of the tests in the battery was a "mental test." The students were given the option of taking the "mental test" in either English or Arabic with the understanding that scores on one form would not be compared with those on the other form. Approximately 70 per cent of the students chose the Arabic test. Of those taking the Arabic Test of Mental Ability, 291 completed the first semester at the university. Of those taking the Henman-Nelson Test of Mental Ability, 96 completed the first semester's work.

During the Spring term, 1953, 82 twelfth-grade students in the International College (affiliated with the American University) were given the vocabulary test. A 30-minute time limit was used.

#### E. RESULTS AND DISCUSSION

The product-moment correlation between the Arabic Test of Mental Ability and first semester average grade was .58, between Henman-Nelson Test of Mental Ability and first semester average grade was .51, between the vocabulary test and first semester average grade was .57. All of these correlations are significant at the .01 level of confidence. It is clear that the test used in each instance was an effective predictor of academic per-

<sup>3</sup>This test was constructed by Mr. Husein Mandil under the supervision of the author.

formance. Indeed, the tests seem to be at least as useful as are similar tests in predicting academic success of American college students.

TABLE 1  
COMPARISON OF ACADEMIC AND TEST PERFORMANCE OF ARAB STUDENTS

Group	University freshmen	University freshmen	Prep seniors
N	219	96	82
Test	Group Test in Arabic	Henman-Nelson Test	Vocabulary Test
Mean test score	62.3	58.0	79.5
Std. dev.	11.9	10.9	7.4
Mean semester grade	66.1	67.9	66.3
Std. dev.	9.6	9.3	9.5
r	.58	.51	.57

If, then, we wish to predict performance in an academic situation of the American type, it would appear that group tests of verbal ability are satisfactory for that purpose. If such performance is our criterion of intelligence, then these tests can be called intelligence tests, even in a non-Western setting. Moreover, the traits which we call "verbal ability" are not simply skills in using the English language. Verbal tests in the Arabic language are effective at predicting academic success when courses are conducted in English, so long as all students meet minimal standards of competence in the English language. An Arabic vocabulary test, given to bilingual Arabs, will effectively discriminate between those who are more and those who are less capable of coping with school subjects taught in English.

#### F. SUMMARY

When confronted with the task of developing intelligence tests for Arab students, we took the position that adaptations of American tests would be at least as effective as "culture-free" tests. Our criterion of success was performance in an American university located in the Near East. We employed tests of the type which have been effective at predicting academic success (measuring academic intelligence) in the United States. The Henman-Nelson Test of Mental Ability (in English), an Arabic Test of Mental Ability modeled after the Henman-Nelson Test, and a multiple-choice test of Arabic vocabulary were used. There was a correlation of from .51 to .57 between scores on each of these tests and first-semester grades. Moreover, the tests in Arabic were as effective in predicting success of Arabs in courses conducted in English as American tests are in predicting academic success in America. The results seem to confirm the thesis that an intelli-



gence test can be adapted for use in different cultures if there is a criterion of intelligence shared by those cultures.

## REFERENCES

1. ANASTASI, A., & FOLEY, J. P. *Differential Psychology*. New York: Macmillan, 1949.
2. LUCAS, C. M. *Survey of the Literature Relating to the Effects of Cultural Background on Aptitude Test Scores*. Princeton: Educational Testing Service, 1953.
3. SPIKER, C. C., & McCANDLESS, B. R. The concept of intelligence and the philosophy of science. *Psychol. Rev.*, 1954, 61, 255-266.
4. TURNBULL, W. W. *Influence of Cultural Background on Predictive Test Scores. Proceedings of 1949 Invitational Conference on Testing Problems*. Princeton: Educational Testing Service, 1950.

*Department of Psychology  
American University of Beirut  
Beirut, Lebanon*



## INTRA-COMMUNICATION AND ATTITUDE: A METHODOLOGICAL NOTE\*

*Department of Sociology, Columbia University*

---

W. GODFREY COBLINER

---

### A. INTRODUCTION

It is a fact that the division of science into branches with neatly defined fields is rapidly vanishing. More and more one finds instances of several disciplines being jointly engaged in the study of a single group of phenomena. Thus, in the understanding of human communication the social sciences can no longer claim a monopoly. Mathematics, physics, biology, neurology, and psychiatry have joined in that effort and are finding their place along with the social sciences in the emerging branch of cybernetics. In addition, and perhaps partly as a result of this coöperation, the uni-causal explanation for phenomena in communication still satisfactory a decade ago in each of the scientific disciplines is now giving ground to the pluri-causal approach. It seems timely therefore to review some recent findings in attitude research in the light of this development.

### B. COMPLEMENTARITY AND LEVELS OF INTEGRATION

The pluri-causal principle or the axiom that *one* phenomenon is caused by the action of *several* agents rather than by that of a single one, testifies to a maturing of any scientific discipline. It reflects versatility, mastery of procedures, and the existence of a body of integrated theorems. Its adoption paves the way for that ultimate goal of knowledge—a unified science.

In physics a long epoch of natural science came to a close when Niels Bohr introduced the concept of complementarity (1) which went far beyond Heisinger's principle of indeterminacy. It is no longer possible, said the atomic scientist, (a) to describe nuclear phenomena independently from the instruments used in their discovery, (b) to understand these very phenomena by a single, unique objective theoretical model. One may need more than one theoretical model to account for a single datum, to grasp its meaning in full. Another eminent atomic physicist, J. R. Oppenheimer, in discussing this issue, suggests that instances of complementarity may perhaps be met with in biological, psychological, and cultural problems (17).

---

\*Received in the Editorial Office on October 23, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

In biology a somewhat parallel development has come about. Here the traditional quest for the single and unique explanation for observed data has been contested in a slightly different fashion. A. B. Novikoff (16) speaks of levels of integration into which a single datum can be placed. "What were wholes on one level," he writes, "are parts on a higher one . . . knowledge of the lower level is necessary for a full understanding of the higher level; yet the unique properties of phenomena at the higher level cannot be predicted, *a priori*, from the laws of the lower level." An example would be the neuron's performance: where there is a central brain its performance is determined by that organ. If the brain is absent as is the case in the lower animals, then the neuron's function follows different, and what is significant, autonomous laws. Previously such reasoning was known under the name of macro- and micro-analysis.

Evidently distinct elements of relativity are contained in both the principle of complementarity and that of levels of integration. To some extent they have been tacitly applied in the study of human communication. One has but to consider divergent methods. The clinical approach is intensive and qualitative, while that of the social scientist at present, quantitative and global. The one is essentially focussing on the personality system as a unit, the other on the social system, to obtain relevant data. And yet, in the last analysis, both disciplines study the same phenomena, though on different levels. A major point where their interests converge is in the study of message interpretation or meaning which is a central issue in communication behavior and mechanics.

### C. COMMUNICATION AND MESSAGE

The framework of communication is studied by mathematicians. According to the mathematical theory of communication the meaning of a message is exclusively determined by the source—the sender—unless the message is subject to distortion along the channel of transmission. By contrast, there is general agreement today that in human communication the meaning—and hence the effect of a message—is also, and in varying degrees, determined by the receiver (19).

When we study communication phenomena from the vantage point of the group or social level of integration, differences in the message interpretation by the receiver are attributed to what is known as rôle effects (i.e., the receiver's age, sex, occupation, education, class, etc.). Thus the message of a political speaker addressed to a mixed audience may be differently understood by say a worker and a banker, and it is argued that this difference



arises from the dissimilar social position (rôle) of the two recipients. Such an approach stresses the property, the quality, the *what* of the message's meaning. It is an existential, agenetic orientation; it ignores the past; it disregards processes that may occur or have occurred within the *receiver* (worker and banker in this case) unless such processes are immediately linked—in time and space—to rôle performance. This scheme of explanation has a simple way of dealing with discrepancies: they are removed with the aid of refined numerical explication (correlations, variance analysis, etc.) and when this is done, unaccounted deviations from the norm, the anticipated behavior, are labeled dysfunction and abnormality, both carrying a derogatory notion. The researcher takes advantage of the unfortunate linguistic coincidence that norm (statistical) and normality stem from the same verbal root. Whatever the merits of his value judgments may be, from the viewpoint of understanding they remain makeshift solutions. In this case abnormality, to use a term suggested by R. K. Merton, is a blanket concept within which a variety of data is buried and therefore remains inaccessible to scientific investigation.

There is, however, an alternative way of dealing with meaning in this context. If the focus is shifted from its treatment as a mere end result, a *what*, to the preceding process, to the *how did it come about*, suddenly new vistas are opened. One begins to search for the forces that initiated, participated, and determined the outcome of the process that attached the specific meaning to the message. Naturally, this is no longer the group level but the individual level of integration, the person as a whole, no longer one of his rôles, is now the unit of exploration. Also when we do that and emphasize our interest in the process instead of in the subsequent result, we elect to order and interpret data from the dynamic aspect which automatically leads to genetic explanations. It is expedient to single out such data by a special term and to call them intra-communication.

What should be done when the social scientist, employing his tools of research geared to the group level of integration, encounters intra-communication phenomena in the course of his work? Is it advisable to leave their investigation to another discipline? The answer is no, under any circumstances, and rightly so. The very essence of Bohr's argument advises not to rely on the single model, however ingenious, to explain the unsolved occurrence. There is no need to fit the data into the logical straitjacket of the unique model, as for instance one does when the catch-all of abnormality is brought into the game. Why not construct another theoretical model and by using both models retain the freedom that comes with choice,

expanded resources, and methods. For example, one may design a model of the personality system to deal with intra-communication data. A more flexible approach of this kind is useful in research operations, since it is not always feasible to obtain the friendly and willing coöperation of another discipline when the data are located on the borderline between two branches of science. Here marginal knowledge will result in the solution of baffling problems and may enable the investigator to make important contributions to the advance of science.

In the following, certain aspects of intra-communication processes will be discussed in reference to experiments reported in the literature.

For this purpose communication is here defined by the writer in a broad sense as *any act or performance, direct or indirect, of one or more persons which affect the perception, the knowledge, the feelings, the emotions, the thinking, the attitudes, and the actions of one or more persons in a conscious or unconscious way regardless of whether the said act is intended so to result or not.*

#### D. THE CASE OF THE SLEEPER EFFECT

In the course of a well-known war time experiment (9) an information film was shown to a large audience of randomly selected U. S. soldiers to establish whether its effects would show "loss or gain in time." Contrary to all expectations based on current theories of learning and other theoretical considerations, an unusual occurrence was noted. Under rigorous experimental conditions, the measured group effect of a few but crucial communication items was greater 9 weeks after the screening of the film than after the short interval of 5 days. A closer inspection of the data indicated that this belated effect was concentrated in the area of opinion change, that is a shift to a pro-British attitude (intended by the motion picture), while the retention of factual data had decreased as expected.

The experimenters called the phenomenon of increased and belated impact, the "sleeper effect," and to this day it has remained an unexplainable datum and has come to be a powerful challenge to all those who are concerned with the result of learning, the effect of indoctrination, propaganda, or advertising, and particularly to those engaged in the study of attitudes.

At the time, the researchers offered three tentative hypotheses to account for the sleeper effect: (a) the substance of the film was retained, the details forgotten; (b) the film was felt by the viewers to be propaganda. Its message became therefore only effective after the source was forgotten; (c) the details of the film were converted into attitudes with the passage of time. None of these three hypotheses was either confirmed or disproved with the

data at hand. No similar experiment dealing with the sleeper effect has been reported since nor has anyone offered any further explanation of the puzzling phenomenon in the theoretical literature.

A closer inspection of these three hypotheses reveals a remarkable communality. All three attempt to account for the sleeper effect on the basis of *endopsychic* events, that is, developments within the person that come about sometime after the stimulus had ceased to operate; the events are therefore in relative independence from external factors. Clearly, this is an explanation on the individual level of integration, it takes the personality system as the unit of research. The experiment, on the other hand, was by nature of its design and other relevant properties geared to the exploration on the group level, the social system. Strangely enough, techniques meant to investigate phenomena on the group level had uncovered data which could no longer be fully explained with the knowledge and procedures of that level. Then, to cope with that problem, the researchers had, without explicitly saying so, followed Bohr's advice and had used another model, i.e., had formulated propositions on happenings of an intra-personal nature.

It is pertinent, therefore, to set forth the implications of their step.

#### E. THE SLEEPER EFFECT AND INTRA-COMMUNICATION

The intra-personal phenomena, i.e., events and sequences within the personality system which lead to the formation of a new attitude, as suggested in the hypotheses on the sleeper effect, belong to the general area of human communication. As suggested above they shall be designated as *processes by which the subject receives, stores, conserves communication items in the "library of the mind,"*<sup>1</sup> and, via some steering arrangement brings them into connection with the judgment function, the value system, and, above all, the effector or motor system which conveys the subject's response, be it symbolic or otherwise, conscious or unconscious, to the outside world.

A phenomenon similar to the sleeper effect is worthy of contemplation. A generation-old story, told in many versions and using a variety of actors, speaks of the effect of a joke upon three guests at a dinner gathering. The first one bursts into laughter immediately upon hearing the joke, the second laughs one hour later, just when leaving the hostesses home, while the third, so the story goes, to the amazement of his colleagues is having a fit of almost unending laughter the following day right in the midst of a business meet-

<sup>1</sup>Credit for this term goes to B. Roueché, writing the *Annals of Medicine* in the *New Yorker*, 1954, 30, 13.

ing. The story is quite instructive. The popular saying is that these two "late laughers" are just dense, a somewhat unsatisfactory explanation for the sleeper effect.

The first problem encountered in going over the details of the sleeper effect experiment is a factual one. Considering the likelihood that already previous to viewing the film most soldiers had ingested an indefinite amount of information and had been exposed to indoctrination similar to that contained in the film, pro-British influence (through press, radio, lectures, discussion, and conversation), it is conceivable that the film acted as a trigger rather than as a cause when some of the soldiers after viewing it had changed their opinion and indicated it on the fifth day after exposure. If so, then what actually took place with the other respondents showing the sleeper effect was this: the amount of indoctrination previously absorbed by them was not yet sufficient to make them respond to the film in the capacity of a trigger mechanism. Figuratively this can be expressed by saying that the reservoir of accumulated indoctrination was not filled up to the brim. It was another event conveying factual or opinion data, not observed by the experimenters, which accomplished this "spilling over" and thus brought about the opinion change—a shift to a pro-British attitude—after exposure to the film.

Even if this assumption could be disproved experimentally, still another possibility remains.

It will be noted that the film, the source or sender of the communication, was not of a monolithic composition. It contained at least three separate elements: (a) visual units; (b) auditory units which were interlocked with the visual units both conveying factual data of a *concrete* nature; and (c) a commentary (sound item) which normally carries the bulk of the tendentious material predominantly couched in generalizing terms and in most cases of more *abstract* character, in short opinion data.

A simple examination suggests that, on the whole, the visual and auditory units can be perceived and absorbed by most viewers and be treated as neutral learning material. There is no need to take sides, no need for any emotional involvement. Not so the commentary material. Its absorption seems to be a more complex and slower process. Here it is suggested to distinguish between two modes of receiving and storing symbolic data of value content: accretive and assimilative learning. In the case of the motion picture we deal with accretive learning when the viewer, previous to exposure, held no opinion or one that was still diffuse and not contrary to that expressed by the film. Storing is therefore easy. However, if the



viewer's previous opinion diverges from that expressed in the film, storing of opinion data encounters obstacles. Counter-forces anchored in the value system and probably linked up with the self-esteem have first to be overcome before the new opinion is crystallized and established. This is assimilative learning. It can be accomplished on either an *emotional* or on a *rational* plane. The emotional plane will dispense with thinking and then a change takes place within a relatively short time. The opinion change is in some degree independent from the substance. As L. S. Kubie puts it: "Emotions impose an automatic value judgment on experience, . . . give experience either a positive or a negative sign . . ." (11).

On the other hand when the opinion change is taking place on the rational plane, thought processes are intervening which determine the direction, the quality, and to some extent, the duration of this conversion. A battle between the new and the old opinion is fought out and while this lasts the person is "in conflict."

A considerable amount of clinical data collected over decades in many life situations suggest that in such a conflict situation the person's attention to inner processes is impaired, the faculty of introspection is in abeyance (20). While the fight goes on he cannot, as it were, be both participant and observer. Hence he cannot communicate to another person (not even to himself) what he "thinks and feels." Externally this withdrawal of inward attention manifests itself in an opinion survey by the subject's switch to the undecided the "don't know" category of respondents. Later on, when the thought operation has run its course and has dissolved the conflict, energies are freed for introspection and the introspective process returns to normal levels. The respondent is again aware of his opinion, can verbalize it and can communicate it, as the writer has discussed it elsewhere (2).

Intra-communication sequences just described seem to offer a possible explanation for the nature of the sleeper effect.

#### F. A TENTATIVE PROPOSITION ON THE SLEEPER EFFECT

1. It is suggested that, after exposure to the film, the value system of respondents opposing the views expressed in that motion picture was put in a state of labile equilibrium following the inception of intra-communication processes.

2. For that portion of respondents for whom the film acted as an effective trigger device and for another portion who dealt with the ingested new opinion data on an emotional plane, the process of opinion change was quick

and could be verbalized on the 5th day after exposure—perhaps sooner—but the experimental data on that possibility are not available.

3. Many of the other respondents who likewise opposed the views affirmed in the film became the victims of a value conflict which was not of their own making and which they tried to solve by thinking it out. While in the grip of this (emotional) conflict they could not verbalize or otherwise communicate what they felt on the subject of this contest. They may have answered "don't know" or "no opinion" on the first interview (which was 5 days after exposure). The subtle but nevertheless important distinction between these alternatives is not yet sufficiently stressed in questionnaires or in oral interviews.

4. Finally it is suggested that those viewers of the film who came to accept the opinion advanced in the motion picture 9 weeks after seeing it (but had not done so 5 days after exposure) were largely recruited from the previous "don't know" category of respondents. Unfortunately, the design of the experiment does not offer the possibility of confirming this assumption. However, supportive evidence can be found in the fact that in the majority of surveys involving repeat interviews, late decision makers usually come from the reservoir of previously undecided, that is neutral respondents.

If these assumptions are proven correct, another hypothesis explaining the sleeper effect can be formulated which integrates inter- and intra-communication sequences.

The sleeper effect is symptomatic of a state of value conflict in the respondents into which they were drawn involuntarily, and which they are solving by thought processes. The passively experienced conflict drains mental energies away from and lowers the capacity for introspection or inward attention. In consequence of the impaired inward attention, both the value clash and the remedial thought process remain unconscious to the respondent, that is he cannot verbalize them or communicate them. Conceptually worded, this means that behind the sleeper effect *there is a mental state, a passively endured situation involving conflict of emotionally loaded forces (values) precipitated by external influence (communication messages) to which the subject responds by a rational mental process (an activity—thinking) and which eventually resolves and terminates the conflict of emotions.*

From the foregoing it becomes clear that thinking plays a central rôle in opinion (attitude) formation and change. Indeed, the rôle of thinking in the phenomenon of the sleeper effect is indicated in hypothesis (c) of the experimenters which speaks of a conversion of details of the film into

attitudes. Presumably the authors suggested that factual data—that is visual and auditory units of the film—were used by the respondents in a process of abstraction to acquire new attitudes. What is known about the nature of thinking and how is it related to attitude and attitude formation?

### G. ATTITUDE AND THINKING

There is a good deal of agreement on the essential properties and functions of thought and thinking among neurophysiologists, neurologists, psychiatrists, and psychologists (10).

K. J. W. Craik, an experimental psychologist and philosopher of science, suggested in 1943 that the main function of thought is to reproduce a model of external reality and of the person's own potential action in response to it (3). By doing that, alternative ways of acting can be reasoned out beforehand instead of being tried out in reality. Thus thought rehearses future action with little expenditure of energy. Half a century earlier, Sigmund Freud expressed the same idea writing as a neurologist (4) and again in 1911 formulated a much more detailed thesis on the basis of two decades of clinical observation. In essence he speaks of thought as "trial or experimental action" (5).

Since attitude is understood to be a "readiness to act," (15) the connection of the triangle—thought-attitude-action—suddenly assumes another significance.

What a person does in thinking is to prepare for action, and it is not far-fetched to suggest that attitude is a midway stage or step between thinking and action proper.

### H. MODES OF THINKING

Thought then has as its arena an imaginary chessboard on which the mental images, the mental representations and symbols of real things and real persons (including the self) are placed, moved, and manipulated with speed and ease. By such a device the thinking person can calculate and foresee, as it were, the outcome of many alternative ways of behavior in reality while he, the actor, stays in the galleries and observes, criticizes, and decides. Once the thinking person clearly conceives and chooses a particular action, a corresponding attitude is crystallized.

How does such a thought process move ahead when it is induced by the impact of pin-pointed communication designed to impose a specific attitude?

Once again the sleeper effect experiment is highly instructive. The communication material contained in the film, as suggested earlier, was composed of interlocking visual and auditory units conveying factual data on

the one hand, abstract opinion data on the other, via the spoken commentary.

The new hypothesis accounting for the sleeper effect discussed above dealt with one form of thinking only, that of reconciling incompatible abstract opinion data—the ones received from outside with those already entrenched in the value system inside. One must realize that this is but one of the many forms of thinking yet insufficiently explored and classified by empirical science.

In the experimenters' hypothesis (*c*) on the sleeper effect another form of thinking is mentioned. They assume a thought performance capable of transforming concrete data into abstract appraisals, of converting knowledge of external facts into (personal) internal opinions. There is reason to believe that this latter type of thinking is a more active and a more autonomous mental performance than the mere reconciliation of like elements described earlier.

Apart from types of thought processes which may perhaps be arranged along a hierarchical scale ranging from passive to active one finds modes of thinking which refer to mental representations (images, symbols, etc.) utilized in mental operations. It is tempting to raise the question whether different modes of thinking can be found together in a given person or whether a single specific mode of thinking is characteristic of the individual.

A good many experiments have been carried out to explore this area. The most convincing and elegant experiment was conducted in 1943 at the Burden Neurological Institute in Great Britain where modes of thinking were shown to be highly correlated with physiological and neurological events. F. Golla, E. L. Hutton, and W. G. Walter (7) succeeded in showing on a sample of 98 subjects that people can be roughly divided into two groups according to their prevailing mode of thinking: (*a*) persons with visual imagery and association who perform their thinking in the form of pictorial mental representations and (*b*) another group which habitually think in audio-kinesthetic symbols; a small third group seemed to show a mixed pattern. There was a high correlation between these two distinct modes of thinking and the kind of breathing (regular or irregular) which is a permanent feature in a person as measured by the plethysmograph. In addition the characteristic way of thinking was highly correlated with the person's profile of the Alpha-Brainwave on the Electro-encephalogram (*EEG*). Independent research on the Alpha-Wave on very considerable numbers of persons all over the world—in the thousands—confirms that the Alpha-Wave in adults is fixed and that each person has a wave specific to him as individually his as his fingerprints. It has also been established that this Alpha-



Wave can be used to ascertain a person's predominant mode of thinking because it discloses with great precision the level of activity in the optical centers of the brain, that site which presumably deals with visual perception and visual imagery (18, 21). W. G. Walter has reported that the results of the 1943 experiment at Burden have been confirmed on a sample of 600 people, but he has not furnished any details or additional information (22). He nowadays divides the population into pictorial and non-pictorial thinkers. No further experiments of this kind have been made since neuro-physiologists and encephalogram specialists are, as a rule, concentrating their study on brain disorders rather than on normal brain functions.

Fortunately for the non-applied sciences one does not need the neuro-physiologist or his highly expensive equipment to explore conscious thinking operations. The well worded questionnaire is a very handy device for that purpose and thus the subject matter is decidedly within the reach of the social scientist.

### I. COMMUNICATION AND MODES OF THINKING

The implication of the experiments carried out at Burden for communication research and for interpersonal relations, group dynamics, and many related fields cannot be fully appreciated at this stage. In all probability the impact of the findings promises to be vast.

At present no further empirical data have been made available to guide us in the appraisal of these findings. Nevertheless a few ideas seem worthy of discussion. They are submitted with reservation even at the risk of sounding speculative at the moment.

Taking the sleeper effect experiment as a point of convenient departure, one might ask what happened to the pictorial thinkers after exposure to the information film. They are likely to have concentrated their attention on the visual units of the motion picture which conveyed factual data. Their mode of thinking made them convert these factual data into attitudes, provided they had not been committed previously to another attitude, in which case a conflict of values similar to the one discussed earlier might have released another chain of thought processes. By contrast, the non-pictorial thinkers preferring verbal ways of mentation were incomparably more prone to focus their attention on the commentary. Hence their thinking, if any, proceeded on a more abstract level. Here then we encounter a new kind of selective perception which has been little explored in experiments. If this is true, there is good reason for considering an overhaul of existing measuring methods in surveys of communication impact. For the pur-

pose of ascertaining the effectiveness of communication, it will no longer seem advisable to rely on the simple count of answers of respondents who are neatly arranged in stratified groups. Modes of thinking make them unequal in many ways and perhaps it will be impossible to arrange these different modes of intra-communication, of which two have been discussed here, on a single hierarchical scale.

Social scientists, communication experts wherever they may be, in public relations, industry, education, and advertising, will seriously have to reconsider their present methods of investigation and dissemination. How effective, for example, is an information film whose visual, auditory, and commentary units—a differentiation suggested above—are closely interlocked when different portions of the audience (as yet unknown) respond to these stimuli separately and in ways that would demand an asymmetrical distribution of the three units? How can the effect of a communication item be measured with currently accepted procedures when complicating phenomena such as the sleeper effect prove to be the rule rather than the exception? Thus the well-established and well-tested methodological procedures of yesterday and today are definitely challenged by the sleeper effect.

It is essential that our knowledge of communication be expanded by opening new territory. What can be learned from the works of art that undoubtedly contain categories of communication as yet unexplored and what can these art works tell us about the rôle of emotions as an effective accessory unit in the dissemination of ideas?

Another point is worthy of consideration. Experiments and careful observation on large groups of children and adults suggest that man is not born with the Alpha-Wave characteristic of his adulthood. The Alpha-Wave pattern emerges in the first years of the second life decade (8, 12, 13, 21, 22). Unfortunately, there are no data available at this time to tell us whether this late establishment of the Alpha-Wave is due merely to a biological maturation of an inherited pattern or whether and to what extent environmental factors play a part. Can the Alpha-Wave pattern be changed through training of a specific kind? Easier to obtain, but also not available, is knowledge on the proportional distribution of Alpha-Wave types among the different sectors of modern society (occupation, class, urban, rural residents, etc.). Data of either kind will give us a clue to the rôle of inheritance, tell us about the impact of education, emotional atmosphere on the formation of Alpha-Wave patterns. The same is true of thinking modes. We do not know whether they are inherited or acquired, whether they can be changed through systematic training. We also lack data on the dis-

tribution of visual and non-visual thinkers among the strata of modern society. Since modes of thinking and the Alpha-Brainwave appear to be related in a symptomatic way (7, 21, 22) a two-pronged approach promises to speed up the gathering of information. There should be: (a) An extensive use of questionnaires to establish the genetic emergence of thinking modes and their eventual distribution in the adult population. (b) A systematic gathering of data on the genetic emergence of the Alpha-Brainwave and its distribution in the adult population. This should be obtained from hospitals, clinics, and other institutions.

#### J. MODES OF THINKING AND COMMUNICABILITY

What is the difference between the pictorial and non-pictorial ways of thinking in regard to awareness and action? Clinical observation seems to indicate that the pictorial way of thinking is limited in that it obstructs the individual's capacity of apprehending and communicating relationships (20). A parallel phenomenon can be observed in the field of linguistics. Language based on vision is in certain respects inferior to verbal language, as the case of gestural language suggests. It renders more difficult the storing of data in the mind. The true social language is indeed verbal language; knowledge is easier shared via its medium. Visual thinkers are by no means inferior to non-visual thinkers; they apparently use additional ways of solving problems of extreme abstract nature such as relations etc., but by and large it seems to be true that they cannot describe to us how they do it.

On the other hand one may assume provisionally that pictorial ways of thinking are closer to action since they bypass the need for abstraction and reconversion into concrete data. Perhaps one will find in experiments that in visual thinkers the interval between an impulse to act and the action proper is much smaller than in the non-visual thinker, since the pictorial thinker sees the planned action before his eyes. Essentially the pictorial thinker deals with the "person," the "thing," the object, while the non-pictorial thinker operates with an attribute of the thing, in the case of words, with the name of the thing. Whether pictorial thinkers therefore tend to form different *structures* of attitudes from those characteristic of non-pictorial thinkers is also worthy of exploration. Perhaps their attitude is of a kind that is "nearer to action," perhaps it consists of a different proportion of cognitive-rational and emotional elements which in turn may affect their stability and susceptibility to change. All this knowledge promises to open new perspectives in attitude research and understanding.

One final point should be mentioned. What is true of information, advertising, and propaganda may also apply to learning. Ways of teaching in

our schools seem to lay heavy emphasis on teaching in an abstract way even in the case of concrete subjects. It is true that abstract ways insure a certain uniform speed in education. But it is quite conceivable that (a) at certain ages a more visual way of teaching might help the child to acquire knowledge more readily, and further that (b) students who are pictorial thinkers are unjustly handicapped by our present prejudice to have one, unique way of teaching instead of considering differential mental make-up in individuals and adjust education accordingly; (c) that perhaps to a certain extent visual and non-visual ways of thinking can be learnt just as any other subject matter.

It is fitting to conclude these notes with a quotation from Sigmund Freud who back in 1923 succeeded, in a statement of utmost conciseness and simplicity, in presenting the entire problem in the following words:

We must not be led away, in the interests of simplification perhaps, into forgetting the importance of the optical memory residues—those of *things* (as opposed to *words*) or to deny that it is possible for thought processes to become conscious through a reversion to visual residues, and that in many people this seems to be the favorite method. . . . We learn that what becomes conscious is as a rule only the concrete subject matter of the thought, and that the relations between the various elements of this subject matter, which is what specifically characterizes thought, cannot be given visual expression. Thinking in pictures is, therefore, only a very incomplete form of becoming conscious. In some way, too, it approximates more closely to unconscious processes than does thinking in words and it is unquestionably older than the latter both ontogenetically and phylogenetically (6).

#### K. SUMMARY

The study of communication is no longer the exclusive concern of the social sciences, since a number of other disciplines with a different focus have joined in that effort. With the growing insight it has now become possible to describe and explain a single phenomenon from various levels, for example, from that of the group level or social system as well as from the individual level or personality system. The approach is illustrated in a discussion of an experiment in mass communication.

The need for a separate exploration of an event in communication and that of the process which brought it about is recommended as a logical consequence. The distinction between communication proper, a phenomenon on the group level encountered between sender and receiver, and intra-communication which is a phenomenon within the receiver, the individual, is suggested. Both are indispensable for a full understanding of communication data.



The place of thinking in intra-communication sequences is discussed in reference to another experiment. Two modes of thinking, visual and non-visual, and their properties are presented.

The eventual need of adjusting current methodological procedures to deal with the new developments is raised and other applications of the findings are briefly mentioned.

## REFERENCES

1. BOHR, N. On the notions of causality and complementarity. *Dialectica*, 1948, 2, 312-319.
2. COBLINER, W. G. On the place of projective tests in opinion and attitude surveys. *Int. J. Opin. & Atti. Res.*, 1951-52, 5, 480-490.
3. CRAIK, K. J. W. *The Nature of Explanation*. Cambridge: Cambridge Univ. Press, 1943.
4. FREUD, S. Project for a Scientific Psychology. (Written in 1895.) In *The Origin of Psychoanalysis*. New York: Basic Books, 1954.
5. ———. Formulations Regarding the Two Principles in Mental Functioning. (Written in 1911.) *Collected Papers IV*. London: Hogarth Press, 1925.
6. ———. *The Ego and the Id*. London: Hogarth Press, 1927.
7. GOLLA, F., HUTTON, M. B., & WALTER, W. G. The objective study of mental imagery. *J. Ment. Sci.*, 1943, 89, 216-222.
8. HENRY, C. E. Electroencephalograms of normal children. *Monog. Soc. Res. Child Devel.*, 1944, No. 3.
9. HOVLAND, C. I., LUMSDAINE, A. A., & SHEFFIELD, F. D. *Experiments on Mass Communication*. Princeton: Princeton Univ. Press, 1949.
10. KUBIE, L. S. Some implications for psychoanalysis of modern concepts of the organization of the brain. *Psychoan. Quart.*, 1953, 22, 21-68.
11. ———. The Place of Emotions in the Feedback Concepts. *Cybernetics*, 9th Conference. (H. V. Foerster, Ed.) New York: Josiah Macy Jr. Foundation, 1953.
12. LINDSLEY, D. B. Electric potentials of the brain in children and adults. *J. Gen. Psych.*, 1938, 19, 285-306.
13. ———. A longitudinal study of the occipital alpha rhythm in normal children: Frequency and amplitude standards. *J. Genet. Psych.*, 1939, 55, 197-213.
14. MACKINNON, D. W. Fact and fancy in personality research. *Am. Psychol.*, 1953, 8, 138-146.
15. MURPHY, G. *Personality*. New York: Harper, 1947.
16. NOVIKOFF, A. B. The concept of integrative levels and biology. *Science*, 1945, 101, 209-215.
17. OPPENHEIMER, J. R. The age of science, 1900-1950. *Sci. Amer.*, 1950, 183, 20-23.
18. OSTOW, M. Psychic function and the electroencephalogram. *Ar. Neur. & Psychiat.*, 1950, 64, 385-400.
19. SHANNON, C. E., & WEAVER, W. *Mathematical Theory of Communication*. Urbana: Univ. Illinois Press, 1949.
20. RAPAPORT, D. *Organization and Pathology of Thought. Selected Sources*. New York: Columbia Univ. Press, 1951.
21. WALTER, W. G. Normal Rhythms: Their Development, Distribution, and Significance. *Electroencephalography*. (D. Hill & G. Parr, Eds.) London & New York: Macdonald & Macmillan, 1950.
22. ———. *The Living Brain*. New York: Norton, 1953.

308 West 106th Street  
New York City 25



PERSONALITIES IN FACES: IV. A DESCRIPTIVE ANALYSIS  
OF THE PERCEPTION OF WOMEN'S FACES  
AND THE IDENTIFICATION OF SOME  
PHYSIOGNOMIC DETERMINANTS\*

*Department of Psychology, Emory University*

PAUL F. SECORD AND JOHN E. MUTHARD<sup>1, 2</sup>

A. INTRODUCTION

In a program of study of the perception of people, two aspects of judgments of photographs might have particular significance: (a) the degree of agreement among judges and the implications thereof; (b) the "disagreement" or variability among judges. The first problem has been investigated and analyzed extensively in a monograph by Secord, Dukes, and Bevan (4). This study grew out of a previous one by the same authors (3) in which an unsuccessful attempt was made to influence the perception of photographs by attaching occupational labels to them. A recent cross-cultural study by Secord and Bevan (2) indicated further that certain aspects of the perception of photographs were quite general, being found among Norwegian and American judges alike. Stritch (6) has augmented the findings of the original monograph by varying physiognomy experimentally. His results likewise reveal considerable agreement among judges and, in general, support the earlier studies.

The second proposition concerning variability among judges has been investigated by Secord and Muthard (5), who demonstrated that individual differences among judges in the perception of women's photographs are significantly related to age and sex of the judge and probably to the personality of the judge. With the exception of this study on individual differences, the investigations cited have employed male photographs.

B. PROBLEM AND PROCEDURE

The present inquiry consists of an intensive analysis of the factors involved in the perception of photographs of women. In general, the analytic and

\*Received in the Editorial Office on October 26, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>The second author is now at the Veterans Administration Hospital, Danville, Illinois.

<sup>2</sup>The writers gratefully acknowledge the financial assistance of the University Center in Georgia in support of the present study.

methodological procedures employed in the monograph by Secord, Dukes, and Bevan (4) are applied. For a detailed account of these procedures, reference is made to the monograph. These operations include the various statistical indices and also the particular use of Tryon's cluster analysis (7) and analysis of variance.

### 1. *Photographs*

In an attempt to acquire more control over the stimulus materials than was accomplished in the previously cited studies, the authors took under standardized conditions two photographs each of about 50 saleswomen. Subjects who appeared to be between the ages of 20-30 were selected. They were instructed to "pose naturally," to refrain from smiling, and to look straight at the camera without turning or inclining their head. The 24 photographs which were technically superior and which appeared to best represent the individual's appearance in life were selected for investigation (see Figure 1).

### 2. *Judges*

Four groups of 20 judges each, from introductory classes in psychology at several universities in Atlanta, Georgia, rated personality, each group rating six different photographs. Three additional groups of 20 judges each rated physiognomy.

### 3. *Rating the Photographs*

Seven-point rating scales similar to those developed for male photographs by Secord, Dukes, and Bevan (4) were constructed by comparable procedures. Physiognomic traits are listed in Table 2. Nineteen of the 35 attributes included in the personality rating scale may be found in Table 1. Instructions for rating personality designated the ends of the scales as "most characteristic" and "least characteristic." Photographs were projected individually on a screen and rated according to standardized instructions (4).

## C. RESULTS

### 1. *Consistency Among Impressions of Physiognomy and Personality*

*Agreement Indices*, indicating the extent to which the various judges agreed on the physiognomic and personality characteristics attributed, were computed for each photograph. The agreement index is an uncorrected Pearsonian  $r$  obtained by splitting the total group of judges into random halves and correlating the resulting pairs of mean ratings (4). Those for physiognomy range from .61 to .93, with a median of .86. Clearly, for all pictures, there is marked agreement among judges on physiognomic traits, al-





FIGURE 1

PHOTOGRAPHS CLUSTERED TOGETHER ON THE BASIS OF PERSONALITY IMPRESSIONS  
 Clusters are indicated by the Roman Numerals I, II, III, and IV. Photographs within a cluster yield similar (but not identical) personality impressions, as determined by a cluster analysis of the personality ratings.

though the degree of agreement varies for different photographs. Similar indices for personality impressions range from .36 to .87, with a median of .59.

## 2. *Eccentricity in Physiognomic and Personality Attributes*

*Eccentricity indices*, representing the extent to which the individuals pictured are rated as possessing attributes in outstanding degree, were computed for each photograph by first converting trait ratings to standard scores. The sigma of the  $Z$  scores for a photograph then constituted the eccentricity index. Physiognomic indices ranged from 12.86 to 6.73, with a median of 9.50. Personality indices ranged from 13.47 to 6.00, with a median of 7.83. Physiognomic eccentricity and personality eccentricity were also shown to be associated by an obtained  $r$  of .50 ( $df = 20$ ;  $P < .02$ ) between the two sets of indices. In other words, individuals whose physiognomic traits were rated toward the extremes tended to have personality traits also rated toward the extremes.

## 3. *Objectivity of the Physiognomic and Personality Attributes*

The "reliability" for various traits was obtained by dividing the judging group in half and computing a Pearsonian  $r$  between the 24 pairs of mean ratings on a given attribute, each pair representing a different picture. No "correction" was applied. Only one physiognomic trait, with an *objectivity index* of .42, *asymmetrical-symmetrical face*, was difficult to rate; others ranged from .75 to .96, with a median of .86.

Personality impressions varied much more, ranging from  $-.31$  to  $+.88$ , with a median of .52. Poor agreement was explained by two factors: (a) failure to understand the meaning of the term; (b) failure to rate attributes pertaining more strictly to behavior. The former is illustrated by frequent requests from the judges for definition of certain words. The second statement is an inference drawn from a consideration of the objectivity indices for the attributes. For example, the traits *aggressive*, *outspoken*, *energetic*, and *reserved*, were low in objectivity, whereas *sexually attractive*, *sincere*, and *feminine*, which appear to represent more intrinsic aspects of the person, were higher in objectivity.

## 4. *Trait Clusters*

An application of Tryon's cluster analysis (7) to the ratings on personality revealed that certain correlated traits clustered together to form more general "traits" or impressions. These clusters are indicated in Table 1. The data

suggest that the judges react to the women in the photographs primarily in terms of what might be called their moral character, social acceptability,

TABLE 1

GROUPING OF IMPRESSIONS ACCORDING TO CLUSTER ANALYSIS  
(Traits are grouped on the basis of an application of Tryon's cluster analysis to a correlation matrix representing the intercorrelations among the traits. Arbitrary cluster-names are assigned.)

Cluster A <i>Moral Character</i>	Cluster B <i>Social Acceptability</i>
Knows all the angles Loose morals Kind* Faithful* Good-mother type* Conscientious* Sincere*	Gay Sense of humor Intelligent Refined Likeable Good mixer
Cluster C <i>Golddigger Syndrome</i>	Cluster D <i>Sexuality</i>
Conceited Likes men's attention Demanding	Sexually passionate Sexually attractive Feminine

Note.—Traits not listed above did not correlate sufficiently with other attributes to be fitted into clusters. Of these remaining traits only two, "easily hurt" and "sharp-tongued" had an objectivity index above .50.

\*Traits marked with an asterisk were reflected; i.e., the algebraic sign of the correlation was reversed.

sexuality, demanding aspects, and possibly in terms of two specific traits, *easily hurt*, and *sharp-tongued*.

##### 5. *Relationships between Impressions of Physiognomy and Personality*

A test of the hypothesis that there was association between impressions of physiognomy and impressions of personality was made by conducting the following two steps: (a) By means of Tryon's cluster analysis, each photograph was placed in a group so that the pictures in any one group had similarly rated *personality* attributes and different groups had dissimilar traits. The 19 photographs in the four clusters yielded by this process are shown in Figure 1. The remaining five photographs did not fit into any clusters and were omitted from this analysis. (b) An analysis of variance was performed to ascertain whether photographs within groups were more similar, with respect to physiognomy, than photographs from different groups. From the obtained *F*-ratio of 2.49 ( $df = 63/315$ ;  $P < .001$ ) for the interaction between clusters and traits, the conclusion may be drawn that, for the four

clusters as a whole, there are combinations of physiognomic traits which are related to patterns of personality impressions.

A further pertinent question concerns the degree to which various aspects of physiognomy and more influential than others in determining judgments of personality. Relevance indices were computed to answer this question. These essentially represent the mean degree of similarity on a given physiognomic trait among each cluster of photographs with similar personalities (see 4, pp. 256-259). The degree of relevance of the various traits may be observed in Table 2. Of significance is the point that most, if not all, of

TABLE 2  
EXTENT TO WHICH SPECIFIC PHYSIOGNOMIC ATTRIBUTES ARE RELATED TO PERSONALITY PATTERNS OF THE CLUSTERS

(The relevance index essentially represents the mean degree of similarity on a given physiognomic trait among each cluster of photographs yielding similar personality impressions.)

Physiognomic trait		Physiognomic trait	
Skin texture	5.60	Mouth (compressed—relaxed)	3.27
Lipstick	5.12	Eyebrows (natural—madeup)	3.20
Grooming	4.92	Facial structure	2.95
Eyes (dark iris—light iris)	4.50	Mouth (smiling—not smiling)	2.92
Lips (thick—thin)	4.35	Eyes (dull—bright)	2.85
Eyebrows (high—low)	4.22	Eyebrows (straight—arched)	2.25
Face (plain—beautiful)	3.92	Lips (bow-shaped—straight)	2.25
Shape of face	3.90	Eyelid visibility	1.92
Eyes (narrowed—widened)	3.82	Age	1.90
Estimated height	3.77	Hairdo	1.30
Estimated weight	3.66	Tilt of head	1.15

the physiognomic characteristics apparently contribute something to personality impressions, as evidenced by the fact that all relevance indices are greater than 0.00.

The specific nature of the association between impressions of physiognomy and of personality are spelled out in Table 3. Here are listed the physiognomic attributes rated towards the extremes of the scale and held in common by those photographs having the highest mean rating on a trait cluster and those having the lowest. Of course, appearance of the physiognomic traits in common among the three most extreme photographs does not guarantee that these traits are determiners of the impression represented by the personality cluster, because of the possibility of some congruence occurring by chance. The previously reported analysis of variance shows, however, that there are significant relationships between physiognomy and personality, and the probability that the physiognomic traits listed in Table 3 are determinants of personality impressions representing the cluster extremes shown is greater than for those not listed.



Important to keep in mind is the fact that attributes found to be common to a set of pictures representing the extremes of a trait cluster are selected from Z-score distributions, so that the reported traits occasionally represent scale extremes only in the sense of relativity to the other pictures.

TABLE 3  
PHYSIOGNOMIC ATTRIBUTES CHARACTERISTIC OF THOSE PHOTOGRAPHS RATED EXTREME IN PERSONALITY TRAIT CLUSTERS

(Physiognomic attributes are listed for photographs averaging  $\pm 1.00S$  in the same direction from the Z-score mean. Remaining traits, marked with an asterisk, permit one photograph to deviate from this criterion, provided the criterion is met by the three pictures at the opposite end of the cluster.)

Trait Cluster	Direction of Rating	
	High (photos CPT)	Low (photos ROU)
<i>Moral Character</i>	Bright eyes Widened eyes	Bow shaped lips
	High (photos IJT)	Low (photos LOU)
<i>Social Acceptability</i>	Eyelids not visible Smooth skin Well groomed Smiling mouth	Unsmiling mouth Poorly groomed*
	High (photos KOR)	Low (photos CPV)
<i>Goldigger Syndrome</i>	High eyebrows Bowed lips Visible eyelids Tilted head Narrowed eyes*	Square face Widened eyes Untilted head Straight lips*
	High (photos EIR)	Low (photos LNP)
<i>Sexuality</i>	Narrowed eyes Relaxed mouth Beautiful Thick lips* Smooth skin* Well-groomed* Much lipstick*	Thin lips Straight lips Coarse skin Poorly groomed Little lipstick Compressed mouth Plain

#### D. DISCUSSION

The present results again confirm the finding of five previous studies (2, 3, 4, 5, 6) that judges show marked agreement when asked to rate photographs. Observations of the rating process and analysis of the ratings leave little doubt that judges presented with this task find it a relatively easy one; furthermore, they make their ratings with a facility which suggests the presence of practiced stereotypes.

Photographs of women, like those of men, vary widely in the "strength" of the impressions of personality they make upon the viewer. Some pictures give rise to many extreme ratings, others to few. This variation in person-

ality eccentricity is associated with physiognomic eccentricity, in the case of female photographs, in a manner comparable to the relationship previously demonstrated for males (4). A study by Stritch (6), using a different set of male pictures, did not confirm this result; however, certain differences in design are believed to disqualify his findings as a crucial test of this particular point.

The median objectivity of personality trait ratings of women's photographs was .52, much lower than the median of .83 for male photographs previously found. Since equal care went into the construction of the rating scales for male and female pictures, the reasons for this difference probably lie elsewhere. Among the possible reasons are: (a) The higher degree of standardization of photographs in the present study. (b) Women are not as readily stereotyped as men, on the basis of their physiognomies. (c) The greater homogeneity of the sample; i.e., all were young women. Present data do not permit a check on these hypotheses.

As might be expected, traits having the highest objectivity are quite different for male and female photographs. Those traits attributed to photographs of women on which observers agreed most closely were sexually attractive, sincere, feminine, likes men's attention, refined, good mixer, and faithful. For photographs of males, the most agreed-upon traits were cheerful appearance, sense of humor, self-confident, intelligent look, determined look, likeable, and honest face. A direct comparison of these results is not possible, however, since preliminary work with the rating scales produced final scales which were most appropriate for male and female pictures, respectively, without many traits common to both.

The present study confirms two earlier ones in demonstrating a significant relationship between physiognomy and personality (4, 6). Physiognomic attributes found to be most potent in determining personality impressions of women were skin texture, amount of lipstick, grooming, darkness of iris, and thickness of lips. Most relevant for men were age, skin texture, thickness of lips, facial tension, horizontal wrinkles in brow, and eye depth. Since the age range for the female photographs was restricted, and for the males not, ratings on age may not be compared here. Skin texture and thickness of lips appear to be important determinants of personality impressions among both sexes. The other traits appear to hold somewhat distinctive places for their respective sexes.

Of the four personality trait clusters, sexuality appears to be most definitely linked to physiognomy. Selection of photographs representing the extremes on sexuality led to the identification of more physiognomic traits

acting as determinants than in the case of any of the other trait clusters. One might tentatively interpret this relationship as a reflection of the stress laid by our culture upon sexuality as a central component of the personalities of young women.

Some comment on the applicability of the present results to the perception of women in real life is perhaps appropriate at this point. The writers do not assume that the perceptions occurring under the conditions of this investigation can be generalized to life situations without further evidence. One might object to the present experiment on the grounds that, under the restricted experimental conditions, judges resort to "logical" bases of judgment which they do not normally use; i.e., a smiling face means sociability or an overuse of lipstick implies sexuality. This argument is rather weak, however, for on what basis are these judgments logical? Such associations cannot be arrived at through formal logic alone; they have some empirical basis in the observations made in our own culture. Thus, it may well be that such associations made in the laboratory may also be made in life situations.

The fact that available cues are much reduced in the experimental situation, leaving fewer for the judge to choose from, is a more cogent argument against generalizing from the present findings. Outside the laboratory added behavioral cues may overshadow physiognomic cues. Only further investigation involving judgments of actual persons can settle the issue.

The present study, considered together with the previous one, on individual differences (5), has clearer implications for application to situations where photographs are employed. The use of the analytical and methodological procedures outlined briefly here and in detail in (4) to develop a projective test similar in purpose to the Szondi test (1) is one example. With their utilization, one would not be dealing with unknown stimulus materials and one could eliminate eccentric photographs which tend to give universally-shared impressions in favor of photographs with low agreement indices which yield different reactions from different individuals.

#### E. SUMMARY

As part of a series of studies concerned with the perception of people 140 college students rated 24 photographs of young women on sets of physiognomic and personality attributes. An intensive analysis revealed the following:

1. Judges agree, to a fairly high extent, in rating physiognomic and personality attributes.

2. Photographs vary in the extent to which they are rated as possessing either physiognomic or personality attributes in outstanding degree. Physiognomic eccentricity correlates with personality eccentricity.

3. Physiognomic traits were rated with a high degree of reliability. Personality attributes on the other hand, varied considerably in the extent of agreement obtained. Many traits were rated with fairly high agreement, but some adjectives in the personality scale were unfamiliar or ambiguous. Furthermore, attributes pertaining more strictly to behavior, e.g., *outspoken*, were generally not rated with high agreement.

4. By applying Tryon's cluster analysis, personality characteristics were found to fall into four groups. Their content is suggested by the cluster names assigned: *moral character*, *social acceptability*, *gold-digger syndrome*, and *sexuality*.

5. By analysis of variance it was determined that there are combinations of physiognomic traits which are related to a significant degree to patterns of personality impressions. Physiognomic traits responsible for extreme personality impressions were identified and the degree to which the various traits contributed to personality impressions ascertained.

6. The above results for female photographs were compared with those for male photographs obtained in earlier studies, and their general implications discussed.

#### REFERENCES

1. RABIN, A. I. The Szondi test. In H. H. Anderson and G. L. Anderson (Eds.) *An Introduction to Projective Techniques*. New York: Prentice-Hall, 1951.
2. SECORD, P. F., & BEVAN, W. Personalities in faces: III. A cross-cultural comparison of impressions of physiognomy and personality in faces. *J. Soc. Psychol.*, (in press).
3. SECORD, P. F., BEVAN, W., & DUKES, W. F. Occupational and physiognomic stereotypes in the perception of photographs. *J. Soc. Psychol.*, 1953, **37**, 261-270.
4. SECORD, P. F., DUKES, W. F., & BEVAN, W. Personalities in faces: I. An experiment in social perceiving. *Genet. Psychol. Monog.*, 1954, **49**, 231-279.
5. SECORD, P. F., & MUTHARD, J. E. Personalities in faces: II. Individual differences in the perception of women's faces. *J. Abn. & Soc. Psychol.*, (in press).
6. STRITCH, T. M. Experimentally induced changes in physiognomy and personality impressions. (Unpublished doctor's dissertation, Emory University, 1954.)
7. TRYON, R. C. *Cluster Analysis*. Ann Arbor: Edwards, 1939.

*Department of Psychology*  
*Emory University*  
*Emory University, Georgia*



## THE SPACE-VISUALIZATION FACTORS RELATED TO TEMPERAMENT TRAITS\*

*Department of Psychology, University of Delaware*

---

ERNEST S. BARRATT

---

### A. PURPOSE

Recent research has indicated a trend to assess achievement and aptitude tests as potential objective measures of the conative and affective aspects of the personality (2, 3, 6, 7, 12). This research has been primarily stimulated by the extended use of factor analysis and pattern analysis in clinical psychodiagnostics (2, 3, 6). Paralleling this trend has been the increasing consideration given to perception as a key intervening variable in various personality theories (8, 13). The present study is contiguous to both trends.

The general purpose of this study is to explore the relationship between the space-visualization factors and certain temperament traits. The results should serve a two-fold purpose: (a) help clarify the nature of the so-called space tests; (b) increase the value of the space tests in counseling and guidance work.

Thurstone (9) isolated three space factors in the area of visual thinking which he tentatively defined as follows:

The first space factor ( $S_1$ ) represents the ability to recognize the identity of an object when it is seen from different angles; . . . the second space factor ( $S_2$ ) represents the ability to imagine the movement or internal displacement among the parts of a configuration that one is thinking about; . . . the third space factor ( $S_3$ ) represents the ability to think about those spatial relations in which the body orientation of the observer is an essential part of the problem . . . one might surmise that kinesthetic imagery is somehow involved in this factor and that kinesthesia constitutes the characteristic of the third space factor (6; pp. 2-3).

Several studies have attempted to define these space factors in terms of item complexity (14) and the problem-solving processes used by the subjects taking the tests (1).<sup>1</sup> It has also been suggested that tests typical of these

---

\*Received in the Editorial Office on October 29, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>For a complete history of the space factor, see (4).

factors correlate significantly with certain temperament traits (11).

The present study was designed to explore these suggestions further by testing the following hypotheses:

1. The central tendency for males and females will be significantly different on the space tests.
2. There will be a difference in the factorial composition of the space tests for male and female; a general space factor will be obtained for the females and three distinct space factors for the males.
3. Rorschach movement responses and reflection (Thurstone Temperament Scale) will be related to Space 3 ability, both being indicative of a high degree of ideational content.

### B. PROCEDURE

Ten space tests, the Thurstone Temperament Schedule, and the Group Rorschach were administered to a population of 199 college students, 103 females and 96 males. The tests were administered under standard conditions and all tests were scored twice by independent scorers.

The Thurstone Temperament Schedule was chosen to measure temperament traits because the same items are not scored in more than one dimension and also because it was constructed to emphasize "important stable traits which describe how normal, well-adjusted people differ from each other" (10, p. 11).

The test variables included:<sup>2</sup>

1. Barratt-Fruchter Chair-Window Test, Form *A* (*CWA*);  $S_3$ ; the task is to determine through which of five windows one would have to look to see a chair from different angles; 32 items;  $r_{11} = .89$ ; 8 minutes.
2. Barratt-Fruchter Chair-Window Test, Form *B* (*CWB*);  $S_3$ ; a modification of *CWA* in which one has to imagine the chair in the room; 32 items;  $r_{11} = .87$ ; 8 minutes.
3. Thurstone's Flag Test (Flag);  $S_1$ ; (Science Research Associates).
4. Thurstone's Figure Test (Fig);  $S_1$ ; (Science Research Associates).
5. Thurstone's Card Test (Card);  $S_1$ ; (Science Research Associates).
6. Minnesota Paper Form Board (*MPFB*);  $S_2$ ; (Psychological Corp.).
7. Differential Aptitude Test (*DAT*);  $S_2$ ; (Psychological Corp.).
8. Industrial Aptitude Survey, Part 10, Spatial Orientation (*PB*);  $S_3$ ; (Psychological Research Center, Los Angeles).

<sup>2</sup>Only those tests which are not sold commercially are described in detail;  $S_1$ ,  $S_2$ ,  $S_3$  after each test refers to the hypothesized factor loading on the spatial tests.

9. Guilford-Zimmerman Aptitude Survey, Part V, Spatial Orientation (*MB*);  $S_3$  or  $S_2$ ; (Sheridan Supply Co.).

10. Guilford-Zimmerman Aptitude Survey, Part VI, Spatial Visualization (*CL*);  $S_3$ ; (Sheridan Supply Co.).

Variables 11 through 17 are the six traits measured by the Thurstone Temperament Schedule (Science Research Associates):

11. Active (*Ac*); constantly on the go; in a hurry.
12. Vigorous (*Vg*); large muscle physical activity.
13. Impulsive (*Imp*); take chances; snap judgments.
14. Dominant (*Do*); a leader and organizer.
15. Stable (*St*); calm, even disposition; not easily distracted.
16. Sociable (*So*); makes friends easily; coöperative.
17. Reflective (*Re*); meditative; reflective thinking; reserved.
18. Movement on Group Rorschach (*M*); weighted according to Zubin's Scale for Levy Movements Blots (15; p. 15-3).

### C. TREATMENT OF DATA AND RESULTS

The means for males vs. females were significantly different on eight of the 10 space tests (Table 1), substantiating the results of other studies (4;

TABLE 1  
A COMPARISON OF MALES AND FEMALES ON SPACE AND TEMPERAMENT TESTS

Test	Range		Mean		$\sigma$		<i>t</i> Ratio
	M	F	M	F	M	F	
1. CWA	26	27	21.4	16.1	5.9	6.1	5.97
2. CWB	27	28	23.3	19.6	5.8	6.5	4.02
3. Flag	61	59	39.8	31.8	12.4	13.3	4.21
4. Fig	52	53	34.1	26.2	12.1	11.5	4.48
5. Card	50	50	28.8	25.8	10.8	10.1	1.94
6. MPFB	38	33	45.0	43.8	8.2	8.1	.980
7. DAT	85	91	66.7	53.4	20.0	22.5	4.26
8. PB	45	46	30.9	24.1	12.1	11.0	3.85
9. MB	50	37	19.6	11.0	10.8	8.1	5.94
10. CL	57	48	26.0	16.9	13.4	11.8	4.75
11. Ac	15	15	10.7	11.0	2.9	2.9	.787
12. Vg	19	17	12.6	7.6	3.8	3.9	8.79
13. Imp	16	16	11.1	11.5	3.6	3.2	.822
14. Do	19	20	10.7	9.7	5.3	4.4	1.38
15. St	17	17	10.6	10.0	3.8	3.6	.924
16. So	19	15	12.5	13.5	3.7	3.3	2.07
17. Re	22	16	8.5	7.0	3.6	3.0	3.14
18. M	31	36	9.2	11.0	5.8	6.2	2.09

p. 392). The variability was practically the same for both distributions. The results substantiate the first hypothesis and indicate the necessity of factor analyzing the male-female matrices separately.

### 1. *Factor Analysis of Male Test Matrix*

The test inter-correlation matrix for males (Table 2) was factor analyzed using the centroid method (5; p. 59); five factors were isolated (Table 3). These factors were rotated orthogonally to isolate the three space factors mentioned above and at the same time approximate a simple structure solution for all factors. The five rotated factors (Table 4) were interpreted as:

Factor I<sub>5</sub>: (*Do, So*)—An extroversion factor; the traits of a typical leader in our culture.

Factor II<sub>4</sub>: (*PB, CL, CWA, CWB*)—A Space 3 factor.

Factor III<sub>4</sub>: (*Imp; Ac, Vg*)—The prototype of masculinity in our culture; taking chances, athletic; *CL* had a relatively high loading on this factor.

Factor IV<sub>8</sub>: (*St, MB; CWA, MPFB*)—The space tests are a combination of Space 2 and Space 3 tests, although the *MB* test had the highest loading; all of these tests involve the visual perceptual reorganization of a rather complex task, and the *MB* test probably also involves kinesthetic imagery. The stability trait depicts calmness and the ability to concentrate. Thus, this factor was interpreted as a perceptual defense factor which would be demonstrated by an individual's remaining calm in a crisis. The *CL* test probably did not have a high loading on this factor because many of the subjects reported that they were using a verbal reasoning process to solve the problems on that test.

Factor V<sub>3</sub>: (*Flag, Fig, Card*)—A Space 1 factor.

### 2. *Factor Analysis of Female Test Matrix*

The test inter-correlation matrix for females (Table 2) was also factored using the centroid method of factor analysis. Four centroid factors were isolated (Table 5) which were rotated orthogonally and defined as (Table 6):

Factor I<sub>4</sub>: (*Do, So*)—An extroversion factor.

Factor II<sub>2</sub>: (*Flag, Fig, Card; MPFB, DAT*)—A Space 1 factor although the Space 2 tests have relatively high loadings.

Factor III<sub>2</sub>: (*Do, Vg, Imp; St, So*)—A fairly general temperament factor, probably a "leadership" factor.

Factor IV<sub>2</sub>: (*CWA, CWB, PB, CL, DAT*)—A Space 3 factor, although also a fairly general space factor.



TABLE 2  
MATRIX CONTAINING TEST INTER-CORRELATIONS (R MATRIX)  
(Decimal Points Omitted)

Above main diagonal are female inter-correlations; below diagonal are male.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. CWA	—																	
2. CWB	669	—																
3. Flag	341	345	—															
4. Fig	373	379	675	—														
5. Card	426	418	722	722	—													
6. MPFB	421	384	387	448	346	—												
7. DAT	428	515	382	368	433	451	—											
8. PB	626	622	424	402	409	405	446	—										
9. MB	517	382	349	487	323	402	342	558	—									
10. CL	579	631	506	532	582	365	580	624	501	—								
11. Ac	116	187	150	153	—014	132	182	171	093	084	—							
12. Vg	271	088	195	163	111	234	102	202	311	249	152	—						
13. Imp	188	109	—082	074	—124	194	038	119	129	063	499	421	—					
14. Do	014	037	005	—015	—081	015	015	036	—009	031	266	183	502	—				
15. St	253	258	211	252	154	314	207	185	329	128	217	174	324	396	—			
16. So	056	067	155	—023	—067	—018	057	—008	—072	—027	301	216	516	697	265	—		
17. Re	039	046	122	085	100	109	172	194	068	162	050	028	041	112	—079	—049	—	
18. M	174	156	—084	—143	046	062	—004	120	086	009	060	053	056	034	089	100	018	—

(Reference for test symbols contained in text under "Procedure.")

TABLE 3  
CENTROID FACTOR LOADINGS FOR MALES

		I <sub>c</sub>	II <sub>c</sub>	III <sub>c</sub>	IV <sub>c</sub>	V <sub>c</sub>	Obt. h <sup>2</sup>	Est. h <sup>2</sup>
1.	CWA	692	199	274	-097	-226	654	669
2.	CWB	670	231	199	-266	-187	648	669
3.	Flag	621	316	-452	080	118	710	722
4.	Fig	635	360	-355	318	103	771	722
5.	Card	587	460	-386	-064	058	713	722
6.	MPFB	583	160	067	228	-085	429	451
7.	DAT	595	272	076	-097	085	450	580
8.	PB	692	292	275	-096	-055	652	626
9.	MB	590	225	117	202	-238	510	558
10.	CL	700	399	099	-165	204	728	631
11.	Ac	371	-371	084	127	148	320	499
12.	Vg	390	-230	101	189	109	263	421
13.	Imp	403	-634	216	220	199	699	516
14.	Do	330	-668	-218	-255	054	671	697
15.	St	458	-268	-112	146	-306	409	396
16.	So	322	-680	-257	-242	092	699	697
17.	Re	159	083	067	-091	235	100	194
18.	M	113	-108	201	-173	-223	144	174

TABLE 4  
ROTATED FACTOR LOADINGS FOR MALES

		I <sub>5</sub>	II <sub>4</sub>	III <sub>4</sub>	IV <sub>8</sub>	V <sub>3</sub>	Rot. h <sup>2</sup>	Orig. Obt. h <sup>2</sup>
1.	CWA	115	665	206	390	110	662	654
2.	CWB	242	694	134	275	155	657	648
3.	Flag	074	235	145	169	778	716	710
4.	Fig	-150	235	225	261	764	780	771
5.	Card	090	394	016	090	746	728	713
6.	MPFB	-101	351	271	385	285	436	429
7.	DAT	078	529	245	100	321	458	450
8.	PB	050	700	268	239	190	657	652
9.	MB	-119	445	168	470	231	514	510
10.	CL	078	674	306	006	425	734	728
11.	Ac	140	-050	504	195	002	314	320
12.	Vg	026	040	459	225	060	267	263
13.	Imp	160	-160	734	268	-163	688	699
14.	Do	669	-245	360	169	012	666	671
15.	St	200	000	178	562	120	401	409
16.	So	669	-286	370	143	045	689	699
17.	Re	035	175	187	-163	100	103	100
18.	M	210	112	087	110	-240	133	144

TABLE 5  
CENTROID FACTOR LOADINGS FOR FEMALES

	$I_c$	$II_c$	$III_c$	$IV_c$	Obt. $h^2$	Est. $h^2$
1. CWA	683	204	-378	186	686	676
2. CWB	629	291	-212	273	600	676
3. Flag	590	327	172	-104	495	612
4. Fig	587	361	185	-228	561	610
5. Card	619	366	135	-307	630	612
6. MPFB	634	233	096	-253	529	593
7. DAT	715	279	089	032	598	590
8. PB	579	280	-204	156	480	509
9. MB	552	261	-155	-076	403	496
10. CL	681	323	076	097	583	590
11. Ac	273	-416	-083	-042	256	405
12. Vg	228	-349	180	181	239	362
13. Imp	392	-561	-182	086	509	481
14. Do	290	-681	-059	-276	628	646
15. St	429	-222	106	101	255	319
16. So	296	-626	-152	-268	574	646
17. Re	125	-084	161	292	134	163
18. M	-034	-083	-151	-210	075	142

TABLE 6  
ROTATED FACTOR LOADINGS FOR FEMALES

	$I_1$	$II_2$	$III_2$	$IV_2$	Rot. $h^2$	Orig. Obt. $h^2$
1. CWA	100	000	-018	822	686	686
2. CWB	-088	088	008	765	494	495
3. Flag	019	519	087	465	494	495
4. Fig	094	605	034	428	559	561
5. Card	180	625	-003	453	628	630
6. MPFB	219	510	092	454	520	529
7. DAT	004	419	157	635	603	598
8. PB	-010	129	-027	675	473	480
9. MB	150	254	-057	555	398	403
10. CL	-072	390	122	640	582	583
11. Ac	319	-154	342	121	257	256
12. Vg	000	-082	480	070	242	239
13. Imp	353	-315	470	261	512	509
14. Do	610	-152	493	-025	639	628
15. St	087	047	415	268	254	255
16. So	620	-182	409	040	586	574
17. Re	-218	-045	270	105	134	134
18. M	249	-047	-100	-060	079	075

## D. DISCUSSION OF RESULTS

Sex differences on the space tests were obvious, the central tendency of the males being significantly higher than the females. The factor composition of the space tests in the female matrix did not show as clear a differentiation into three types of space tests as did the male analysis; the female space factors represented more of a general spatial ability. In neither male nor female analyses was a separate Space 2 factor isolated. When  $S_1$  tests were high,  $S_3$  tests were low, and vice-versa, with  $S_2$  being ranked in between in terms of factor loadings. The factor composition of the space tests could be interpreted as following Zimmerman's hypothesis that space tests are on a continuum from perceptual speed through reasoning in terms of item complexity (14). These data cannot be interpreted as substantiating his hypothesis, but his hypothesis does seem to offer a very logical basis for interpreting the results.

Space and temperament tests did not show much affinity for each other except on Factor  $IV_8$  in the factor analysis of the male data; this factor was defined as a perceptual defense factor. One would expect that ability to remain calm in a crisis would be partially a function of one's ability to organize perceptually a rather complex field. The *MB* test, which had the highest loading of any space test on this factor, involves not only organizing a complex perceptual field, but also more kinesthetic imagery than any of the other space tests included in this battery. Thus, kinesthetic imagery and ability to organize a complex perceptual field probably relate to one's ability to remain calm in a crisis and concentrate; as Thurstone notes, persons who have high stable scores remain calm in a crisis and "claim they can disregard distractions while studying" (10; pp. 1-2).

Reflection and Rorschach *M* responses did not relate to the  $S_3$  tests as had been hypothesized.

Considering only the Thurstone Temperament Schedule, the analysis of the male data produced three temperament factors which were defined as extroversion, masculinity, and stability. The analysis of the female data produced two temperament factors, an extroversion factor and a fairly general factor probably best interpreted as a leadership factor.

## E. SUMMARY

The present investigation was designed to explore the relationship between spatial ability and temperament traits. Only one factor was isolated that included both space and temperament variables; this occurred in the



analysis of the male data and was interpreted as a perceptual defense factor. The factor composition for the male and female analyses was also noted to be different.

## REFERENCES

1. BARRATT, E. S. An analysis of verbal reports of solving spatial problems as an aid in defining spatial factors. *J. of Psychol.*, 1953, **36**, 17-25.
2. CATTELL, R. B. Description and Measurement of Personality. New York: World Book, 1946.
3. FRENCH, J. W. The Description of Personality Measures in Terms of Rotated Factors. Princeton: Educational Testing Service, 1953.
4. FRUCHTER, B. Measurement of spatial abilities: History and background. *Educ. & Psychol. Meas.*, 1954, **14**, 387-395.
5. ———. Introduction to Factor Analysis. New York: Van Nostrand, 1954.
6. JASTAK, J. A plan for the objective measurement of character. *J. Clin. Psychol.*, 1948, **4**, 170-178.
7. PEMBERTON, C. The closure factors related to temperament. *J. Personal.*, 1952, **21**, 159-175.
8. RAMSEY, G. V., & BLAKE, R. R. Perception: An Approach to Personality. New York: Ronald, 1951.
9. THURSTONE, L. L. Some Primary Abilities in Visual Thinking. Psychometric Lab. Report No. 5. Chicago: Univ. Chicago Press, 1950.
10. ———. Examiner's Manual for the Thurstone Temperament Schedule. Chicago: Science Research Associates, 1950.
11. TOMKINS, S. S. Personality and intelligence: Integration of projective and psychometric techniques. In (Hoch and Zubin, Eds.) *Relation of Psychological Tests to Psychiatry*. New York: Grune & Stratton, 1952.
12. WECHSLER, D. Cognitive, conative, and non-intellective intelligence. *Amer. Psychol.*, 1950, **5**, 78-83.
13. WITKIN, H. A., et al. Personality Through Perception. New York: Harper, 1954.
14. ZIMMERMAN, W. S. The influence of item complexity upon the factor composition of a spatial visualization test. *Educ. & Psychol. Meas.*, 1954, **14**, 106-119.
15. ZUBIN, J. Quantitative Techniques and Methods. New York: Columbia Univ. Bookstore, 1950.

Department of Psychology  
University of Delaware  
Newark, Delaware



## GRADUAL STRENGTHENING OF S-R CONNECTIONS OR INCREASING NUMBER OF S-R CONNECTIONS\*<sup>1</sup>

Department of Psychology, Yale University

VIRGINIA W. VOEKS

### A. THE PROBLEM

In the typical learning experiment, only some facet of a series of situations is singled out for notice. Gradual increments often occur in the percentage of occasions this facet of the stimulus-situations is followed by some noted facet of the response-patterns. Classical learning curves show such increments.

The fact that the noted piece of the stimulating situations is followed with increasing frequency by some particular response often is assumed to mean that that *S* is evoking that *R* increasingly often. This assumption, in turn, leads to the idea that an increasing "probability of response" is due to a gradual strengthening of "the" S-R connection. However, gradual increments in  $P_{R+}$  may be due instead to other factors, and not to gradual strengthening of any S-R connection.

For example: (a) In addition to the stimuli we customarily note, many other stimuli are present. These stimuli too may be partial determiners of what response is elicited. Further, these stimuli may differ appreciably from trial to trial. When many different stimuli must become cues for  $R_+$ , and not all of them are present on any one learning trial,  $P_{R+}$  will increase gradually with  $N$  as more and more stimuli become cues for  $R_+$  (cf. Postulate 3 of Guthrie's theory—Voeks 1950). Under such circumstances, the  $P_{R+}$  should increase gradually even though each of the S-R connections is established fully with one pairing of that *S* and *R*. (b) When the on-trial and between-trial stimuli are highly similar (as is usually the case), there are consequent opportunities between trials for extinction of some desired S-R connections, and  $P_{R+}$  can increase only gradually. (c) When different individuals first make the desired *R* on different trials, grouping the data will cause  $P_{R+}$  to increase gradually with  $N$ —even though every subject estab-

\*Received in the Editorial Office on November 9, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This experiment is part of a study presented to the Psychology Department of Yale University, in partial fulfillment of the requirements for the Ph.D. degree. The author is now with San Diego State College.

lished the desired  $S$ - $R$  with full strength when that  $S$ - $R$  pairing first occurred for him.

The three circumstances listed above may be among the primary causes of gradualness in the  $P_{R+}$  increments. This proposition is supported by data previously reported (9). The experiment here reported furnishes additional data bearing on the first two possibilities. It also tests Theorems 7 and 8 deduced (8) from Guthrie's postulates of one-trial learning, one-trial unlearning, postremity, and response-probability being proportional to the percentage of all present stimuli which are cues for that response.

Theorem 7 states: In any specified number of trials after the first  $R_+$ , more alternations in response will occur on the average for subjects in a relatively less stable situation than for subjects in any relatively more stable situation—when (a) not more than half the total stimulus-pattern after the first  $R_+$  is consistently composed of stimuli which are not cues for  $R_+$  and (b) the response between trials is  $R_+$ .

Theorem 8 states: In any specified number of trials following the first  $R_+$ , longer consecutive series of  $R_+$ 's will be made on the average by subjects under relatively more stable stimulus-conditions than by subjects under relatively less stable conditions.<sup>2</sup>

## B. METHOD

### 1. Subjects and Apparatus

The 105  $S$ s were male undergraduates of Yale University. None had studied conditioning. Twenty-five  $S$ s were used in preliminary tests to ascertain that no on-trial stimulus evoked  $R_+$  before conditioning; 32  $S$ s in experimental Group 1; 12 in each of four other experimental groups. The  $S$ s were assigned to the five experimental groups and run in a predetermined random order.

The apparatus is described elsewhere (9).

### 2. Procedures

Various conditioning procedures were used. For all groups,  $R_+$  was a partial or complete blink or wink; the  $US$ , an air-puff;<sup>3</sup> and one of the main

<sup>2</sup>The following definitions were involved in the deductions of the above theorems:  $R_+$  is the response  $E$  is interested in cuing to the on-trial stimulus-patterns; and  $R_i$  is any response incompatible with  $R_+$ . An *alternation* is any occasion when  $R_i$  is made on some trial and  $R_+$  is made on the next, or when  $R_+$  is made on some trial and  $R_i$  is made on the next. A *relatively stable situation* is, roughly, a set of trials which, compared to some other series of trials has (a) greater similarity of the stimuli from trial to trial, and (b) greater dissimilarity of the on-trial and between-trial stimuli. For more complete definitions, see 8, pp. 357-358.

<sup>3</sup>The air-puff was caused by the fall of 100-130 mm. of mercury, magnetically re-



CS's, a markedly damped buzz beginning about .45 sec. before the puff and ending about .09 sec. after onset of the puff. The inter-trial times were 45, 50, 55 . . . 75 sec., arranged in the same predetermined random order for all Ss.

Five experimental conditions were designed to have progressively decreasing stimulus-stability. Decreasing stimulus-stability was accomplished by progressively dropping controls used for Group 1 to reduce variations in on-trial stimuli and to reduce similarity of on-trial and between-trial stimuli.

a. *Procedure 1 (greatest stimulus-stability)*. Each trial began with S breathing in rhythm to E's slowly saying "Ready, inhale, exhale, inhale, hold." On the last "inhale," S inhaled deeply. In response to "hold," S held his breath and "with as short a reaction time as possible" depressed a stiff telegraph key with each hand. This was done in a manner described to S in detail and practiced before the trials began. Depression of both keys activated the buzz. When the buzz terminated, S released the keys and ceased holding his breath. The S was told that this was "a study of individual differences in reaction-time," a statement made highly plausible by the procedure and apparently accepted by all Ss.

Additional controls were introduced to minimize stimulus-variations arising from autonomic changes during the experiment, from differences in what the Ss were thinking about, from changes in their set, from fluctuations in other internal stimuli, and from variations in the external stimuli. A more adequate description of these controls is published elsewhere (9).

b. *Procedure 2*. The key-response was dropped, less effort was made to create and maintain a constant state of calmness in S, and the stated purpose of the experiment was changed. These Ss were told, "This is an experiment in reactive sensitivity. . . ." Far less plausible than the stated purpose for Group 1, this resulted in questions after the session concerning the purpose of the experiment. Probably these Ss had questions *during* the experiment also, which could introduce considerable stimulus-variability from changes in set.

c. *Procedure 3*. The breathing routine was dropped, though the key-response was used. Trials began by E saying "Ready," and after 1 to 7 seconds, "now." In response to "now," S depressed the telegraph keys. The other sources of stimulus-instability present under Procedure 2 were present here.

Deleting the breathing routine decreased stimulus-stability in three ways: released. Pressure of the puff was adjusted to each S's sensitivity and elicited practically a full wink on each trial.

(a) The highly constant and extensive pattern of on-trial stimuli arising from holding one's breath after a deep inhalation were eliminated. (b) The on-trial and between-trial stimuli no longer were as dissimilar. (c) The key-response, though included, no longer could create very stable stimuli. Without the breathing routine and with no other stable pattern of cues substituted, a more variable pattern of stimuli evoked the key-response. Since a response is a function of the stimuli to which it is made, when the stimuli evoking the key-response vary, the key-response also varies. Consequently, *S* made the key-response in a considerably more variable fashion than under Procedure 1 and the stimuli from it no longer would be very stable from trial to trial.

Whereas eliminating the breathing routine unstabilized somewhat the key-response and consequently the stimuli arising from it (in Procedure 3), eliminating the key-response did not unstabilize the stimuli arising from the breathing routine (in Procedure 2). Mainly because of this fact, Group 3 should have had somewhat less stimulus-stability than Group 2. Both groups, of course, would have had considerably less stimulus-stability than Group 1.

d. *Procedure 4.* Neither the key-response nor the breathing routine were used. In addition, the door to the soundproof room was opened; *S* talked with *E* between trials, and *E* moved audibly between and during trials, occasionally rustling papers and shuffling her feet, thus introducing other adventitious variations in the on-trial stimuli.

e. *Procedure 5.* All variations in stimuli present under Procedure 4 were present here, plus some other extraneous stimuli varying from trial to trial—a faint light, a soft bell, and/or a click. One or two of these stimuli were presented on various trials, with varying time relationships to the buzz. No effort was made here to have the *Ss* in a constant state of calmness.

### 3. *Number of Trials*

The mean number of total trials and the mean number of trials following each *S*'s first  $R_+$  are given in Table 1. Both should be equated for the various groups. However, before seeing the photographic records of the response, *E* could not always tell whether a deflection of the light from the eyelid-lever had arisen from a response properly classified as  $R_+$  or whether it had resulted from a shift in gaze or some other incidental response. Therefore the numbers of trials following the first  $R_+$  are not identical for the various groups, but there are no significant differences in either the means or the variances and no systematic differences from group to group.

Although the number of trials was small to avoid (for Group 1) increas-

TABLE 1  
NUMBER OF TRIALS

Group	Mean number of total trials	Mean number of trials after and including the first $R_+$
1	17.2	13.5
2	14.8	12.8
3	15.2	12.2
4	14.8	13.1
5	15.6	12.8

ing fatigue and resultant stimulus-instability, enough trials were given to insure learning: every  $S$  made  $R_+$  on over 20 per cent of his trials.

### C. RESULTS AND DISCUSSION

#### 1. *Tests of Theorem 7: Response-Alternations*

For each individual was computed the ratio of his actual alternations in response to his total opportunities for alternations after the first  $R_+$ . The mean for each group was computed, weighting each subject proportionally to the number of trials upon which his ratio was based (cf. 6, p. 66; 7, p. 21).

Arranging the groups in order with respect to relative stimulus-stability also arranges the groups in order with respect to mean frequency of alternations (see Table 2, Column 3). The order of the five groups is that pre-

TABLE 2  
GROUP MEANS PERTINENT TO THEOREMS 7 AND 8

Groups from relatively stable to least stable stimulus conditions	No. of $S$ s	Mean ratio of actual alternations to maximum alternations possible	Mean ratio of actual length of longest $R_+$ series to longest $R_+$ series possible
1—Most stable	32	.12	.76
2—No key-response	12	.34	.46
3—No "breathing"	12	.42	.45
4—Neither	12	.43	.39
5—Neither, plus other varying stimuli	12	.51	.29

dicted from the theory, a result significant beyond the 1 per cent level of confidence.<sup>4</sup>

<sup>4</sup>There are 120 permutations of five things taken five at a time. Only one of these is the particular order demanded by the theory. Therefore by "chance," one time in 120 five groups would be arranged with respect to some characteristic in the one particular order specified by the theory.

Means may be determined largely by a few deviant individuals. Therefore the means of groups can fit a theory even when the majority of individuals give results at variance with the theory. To check upon this possibility, a Chi-square test was run. This test shows that the number of individuals with relatively few or many alternations is closely related to the degree of stimulus-stability. (The  $\chi^2_c$  is significant far beyond the 1 per cent level of confidence; see Table 3.) In the group with greatest stimulus-sta-

TABLE 3  
RELATIONSHIP BETWEEN STABILITY OF STIMULUS CONDITIONS AND FREEDOM FROM ALTERNATIONS

Percentage of trials after first $R_+$ on which an alternation occurred		Degree of stimulus-stability			Total
		Least stable (Groups 4 & 5)	Intermediate (Groups 2 & 3)	Most stable (Group 1)	
0-15%	$f_o$	1	4	23	28
	$f_c$	(8.40)	(8.40)	(11.20)	
16-45	$f_o$	8	11	8	27
	$f_c$	(8.10)	(8.10)	(10.80)	
46-80	$f_o$	15	9	1	25
	$f_c$	(7.50)	(7.50)	(10.00)	
Totals		24	24	32	80

$$\chi^2_c = 34.0; 4 \text{ degrees of freedom; } p < .01.$$

bility, half of the individuals had no alternations whatsoever after their first  $R_+$ . Such records were made by only 8 per cent of the subjects having less stimulus-stability.

## 2. Tests of Theorem 8: Longest $R_+$ Series

In any specified number of trials following the first  $R_+$ , individuals under progressively more stable stimulus-conditions should have, on the average, progressively more  $R_+$ 's in their longest  $R_+$  series. Data would not necessarily accord with this theorem merely because they accorded with Theorem 7.<sup>5</sup>

<sup>5</sup>Individuals with relatively few alternations do not necessarily have longer consecutive series of  $R_+$ 's than do individuals with more alternations. For example, six  $S$ 's might have 1, 2, 3, 4, 5, and 6 alternations respectively in 10 trials following their first  $R_+$ . Each of these six  $S$ 's could have the same number of  $R_+$ 's in his longest series and have the same percentage of his trials comprised of his longest series: Each might have only two  $R_+$ 's in his longest series, or each might have three  $R_+$ 's, or each might have four  $R_+$ 's in his longest series during the 10 trials—despite the fact that their numbers of alternations were very different in those 10



To test Theorem 8, the number of  $R_+$ 's in each individual's longest unbroken series was divided by his total number of trials after and including his first  $R_+$  (as is demanded by assumptions in the deduction of the theorem). Means for the five groups were computed, weighting individuals proportionally to the number of trials on which their ratios were based.

Each group with relatively more stable stimulus-conditions averaged more  $R_+$ 's in their longest series than any group with less stable stimulus-conditions. (See Table 2, Column +.) The group with the most stable stimulus-conditions had single  $R_+$  series which averaged 76 per cent as long as was possible (i.e., on the average, 76 per cent of the total trials following each individual's first  $R_+$  was comprised of a single unbroken  $R_+$  Series). Groups with progressively less stable stimulus-conditions fell progressively far below this. The order of the five groups is exactly that predicted by Theorem 8, a result significant beyond the 1 per cent level.<sup>6</sup>

Here again, the data for individuals also is that expected on the basis of Guthrie's theory. The greater the degree of stimulus-stability imposed by the experimental design, the greater the number of individuals with relatively long  $R_+$  series. Chi-square is significant far beyond the 1 per cent level of confidence (see Table 4).

### 3. *The Rôle of Internal Stimuli in Conditioning*

For Group 1 vs. Groups 2 and 3, the experimentally controlled conditions differed only in the extent to which various *internal* stimuli were stabilized. However, the resultant response-sequences differed greatly. For example, Group 1 had only about one-third as many alternations as either Group 2 or Group 3. These data demonstrate that internal stimuli (emotional, proprioceptive, and others) can be extremely important determiners of the individual's response, with their rôle not being limited to a motivational one.

trials. Data which accorded with Theorem 7 could be at wide variance with Theorem 8.

<sup>6</sup>The *relative* sizes of the means (not their absolute sizes nor the size of the differences) is the proper test of the theory, since it is only the former result about which the theory made predictions. However, as a matter of interest the statistical significance of the differences in means follow. (a) For number of alternations, using a *t*-test (6, p. 72ff.),  $p < .01$  for Group 1 vs. each of the other groups;  $p \cong .02$  for Group 2 vs. Group 5;  $p \cong .07$  for Group 3 vs. Group 5;  $p > .10$  for Groups 2 or 3 vs. Group 4, and for Group 2 vs. Group 3. (b) For percentage of trials comprised of an unbroken  $R_+$  series,  $p < .01$  for Group 1 vs. Group 4, and for Group 1 vs. Group 5;  $p = .02$  for Group 1 vs. Group 3;  $p = .03$  for Group 1 vs. Group 2;

$p > .10$  for the other comparisons, with  $SE_{diff} = \sqrt{\frac{P_T Q_T}{N_1} + \frac{P_T Q_T}{N_2}}$

TABLE 4  
RELATIONSHIP BETWEEN STIMULUS-STABILITY AND LENGTH OF LONGEST  $R_+$  SERIES

Percentage of trials after first $R_+$ which was com- prised of long- est $R_+$ series		Degree of stimulus-stability			Total
		Least stable (Groups 4 & 5)	Intermediate (Groups 2 & 3)	Most stable (Group 1)	
60-100%	$f_o$	4	4	27	35
	$f_c$	(10.50)	(10.50)	(14.00)	
35- 89	$f_o$	6	13	3	22
	$f_c$	( 6.60)	( 6.60)	( 8.80)	
10- 34	$f_o$	14	7	2	23
	$f_c$	( 6.90)	( 6.90)	( 9.20)	
Totals		24	24	32	80

$$\chi^2_c = 37.7; 4 \text{ degrees of freedom; } p < .01.$$

Some additional empirical evidence is afforded by notes made during the experimental sessions. These notes indicate many instances on which a change in eyelid response coincided with a change in  $S$ 's internal stimulus-patterns.<sup>7</sup> The following are illustrative.

One subject, arms on the table, made 11 successive  $R_+$ 's. He then moved his arms from the table and crossed them; on the next trial,  $R_+$  did not occur. Before the next trial, he placed his arms on the table as before, and  $R_+$  again was made. Another  $S$  made a long series of  $R_+$ 's. He then coughed, and changed his position to such an extent that the beam of light from the eyelid-lever no longer fell on the camera. The  $S$  seemed very disturbed by this, offered rather profuse apologies, and announced he was going to hold his breath for a while so he would not cough again. For the next few trials, no  $R_+$  occurred.

Another  $S$  was extraordinarily restless throughout the experiment, sighed often, shifted his position almost every trial, twice so greatly that the lever became untaped from his eyelid. (This is the only case where that happened.) At the end of the session,  $S$  reported that he had had a recurrent and intense cramp in his leg during the experiment. This  $S$  had an alternation on 73 per cent of his trials, almost twice as many as any other member of that experimental group and six times as many as the mean for that group.

The preceding observations, while far from conclusive, are offered for their heuristic value. Such observations of "incidental" stimuli may furnish helpful additional bases for predicting individual responses.

<sup>7</sup>Since at the time of the experiment,  $E$  was only casually interested in this phase of the problem, the notes are suggestive but lack the completeness necessary for statistical analysis.

#### 4. *Interpretation of the Failure of Responses to Appear Sometimes After Having Once Been Evoked*

Often a response fails to appear after having been evoked on previous trials under similar circumstances. These failures of response are accounted for in different ways by Hull and Guthrie.

Hull would say the response-failures are due partly to (a) variations in the stimulus-patterns present on various trials, (b) differences in the  $s\bar{E}_R$  arising from increments in the various inhibitory potentials during and between trials, (c) "spontaneous" oscillation of  $sE_R$ , and (d) the gradualness with which each  $S$ - $R$  connection is strengthened with repeated reinforcements (3, 4, 5).

Guthrie postulates that each  $S$ - $R$  connection is established with full strength after the response once has been made in the presence of those stimuli (1, 2). According to this theory, the failures of a new response to appear after having been evoked on some previous trial are due solely to variations in the on-trial stimuli and to unlearning between trials of some (or all) of the desired  $S$ - $R$  connections. This unlearning, in turn, is a function of on-trial and between-trial similarity: The greater the similarity of the on-trial and the between-trial stimuli, the greater the number of desired  $S$ - $R$  connections which can be unlearned between trials.<sup>8</sup>

To account for the data of this experiment, we need assume neither gradual strengthening of  $S$ - $R$  connections nor the existence of  $sO_R$ . The Guthrie explanation fits the data nicely: The less the on-trial stimuli differed from trial to trial and the less they resembled the between-trial stimuli, the more sharply curtailed were the response-failures and the less gradual were the  $P_{R+}$  increments. Response-failures virtually disappeared when variations in on-trial stimuli and similarity of on-trial and between-trial stimuli were minimal. No longer did  $P_{R+}$  increase gradually. Instead  $P_{R+}$  increased suddenly after the first  $R_+$  to a near 100 per cent level.

Clearly gradualness of these  $P_{R+}$  increments is related closely to the necessity for forming different  $S$ - $R$  connections on different trials and to the opportunities for unlearning desired  $S$ - $R$  connections between trials.

These data raise some doubts whether  $sO_R$  exists and whether any  $S$ - $R$  connection gains strength gradually. Perhaps all occasions which seem to involve  $sO_R$  and gradual strengthening of  $S$ - $R$  connections involve instead

<sup>8</sup>Increasing the extent to which between-trial stimuli duplicate on-trial stimuli increases also the number of on-trial stimuli the subject must learn to ignore before  $R_+$  can appear with high consistency on the various trials. As a consequence of this too,  $S$ - $R$  connections may appear to be gradually strengthened.

fluctuations in on-trial stimuli and duplication of on-trial and between-trial stimuli.

Do we have any clear evidence that  $sO_R$  exists or that an  $S$ - $R$  connection ever gains strength gradually?

#### D. SUMMARY

From Guthrie's postulates of one-trial learning,<sup>1</sup> one-trial unlearning, postremity, and the postulate that response-probability is proportional to the percentage of all present stimuli which are cues for  $R_+$ , the following theorems have been formally deduced: In progressively more stable situations, there will be on the average (a) progressively fewer alternations in response, and (b) progressively higher percentages of the trials comprised of a single  $R_+$  series. The limiting case is for perfectly stable situations, in which should occur jumpwise curves containing no alternations in response and indefinitely long  $R_+$  sequences.

To test these theorems, conditioned eyelid responses were established for five groups of  $S$ s under progressively more stable stimulus-conditions. That is, each group had progressively less variability in their on-trial stimuli and progressively less similarity of their on-trial and between-trial stimuli.

As stimulus-stability increased, the proportion of trials comprised of a single  $R_+$  series increased and the number of response-alternations decreased to the point where half the  $S$ s had no alternations whatsoever after their first  $R_+$ . Since the data for the five groups fit both theorems, they support the validity of the four Guthrie postulates from which the theorems were deduced.

The data also suggest: (a) On-trial responses are a function of many stimuli besides "the"  $CS$  and  $US$ . (b) Internal stimuli (emotional, proprioceptive, and others) are important determiners of the individual's responses. Were these "incidental" stimuli noted more closely, possibly we would not need to postulate  $sO_R$  nor gradual strengthening of any  $S$ - $R$  connection. (c) Relative similarity of complex situations can be estimated from an *a priori* analysis of the situations, with such estimations furnishing an adequate basis for prediction of responses. Inherent in all theories of behavior is a need for such estimations. (d) Increasing response-probability may be due rarely, or even never, to some reinforced  $S$ - $R$  connection gradually gaining strength. Instead the gradualness with which  $P_{R_+}$  sometimes increases may be due largely to a gradually increasing number of stimuli becoming cues for the desired responses.



## REFERENCES

1. GUTHRIE, E. R. *Psychology of Learning*. New York: Harper, 1935.
2. ———. *Psychology of Learning*. (2nd Ed.) New York: Harper, 1952.
3. HULL, C. L. *Principles of Behavior*. New York: Appleton-Century, 1943.
4. ———. *Essentials of Behavior*. New Haven: Yale Univ. Press, 1951.
5. ———. *A Behavior System*. New Haven: Yale Univ. Press, 1952.
6. LINDQUIST, E. F. *Statistical Analysis in Educational Research*. Cambridge: Riverside Press, 1940.
7. SNEDECOR, G. W. *Statistical Methods*. (2nd Ed.) Ames: Iowa State Coll. Press, 1940.
8. VOEKS, V. W. Formalization and clarification of a theory of learning. *J. of Psychol.*, 1950, **30**, 341-363.
9. ———. Acquisition of S-R connections: A test of Hull's and Guthrie's theories. *J. Exp. Psychol.*, 1954, **47**, 137-147.

*Department of Psychology and Philosophy*  
*San Diego State College*  
*San Diego 15, California*



## MASS ACTION EFFECTS IN LEARNING\*

*Department of Psychology, University of California*

A. ROBERT ROLLIN<sup>1</sup>

### A. INTRODUCTION

Lashley, in 1929, presented the conclusions derived from several years of experimental work on brain damage in rats (19) and advanced a theory of cortical functioning which has since gained wide acceptance. In this paper two important aspects of his theory will be examined. The first is its postulation of a significant non-sensory function of sensory cortex particularly important in complex learning. The second is the implication that transcortical connections are important for complex behavior.

Direct evidence for the non-sensory function of sensory cortex comes from experiments which compare the effects of peripheral and central sensory deprivation upon learning and retention. In summarizing the results of several of these experiments embracing various sensory modalities Lashley wrote: "In all these instances the destruction of the sensory projection area produced a disturbance of function from 9 to 60 times as great as did the destruction of the corresponding peripheral mechanism" (20, p. 8). Studies of this type include Lashley 1929 (19), 1931 (21), 1935 (23), and Tsang 1934 (44), 1936 (45). The above experiments were critically reviewed by Finley in 1941 (13). She concluded that none of them could be considered decisive and, in turn, presented data she felt was inconsistent with the hypothesis of a nonvisual function of visual projection cortex.

With respect to the second point, that of the importance of transcortical connections in complex behavior, Lashley's theory, with its concepts of *equipotentiality* and *facilitative function*, implies an intricate pattern of cortical interaction. A system of connecting fibers lying wholly within the cortex would offer a direct means for such interaction. There is little consistent evidence concerning the importance of transcortical connections in brain function. Certain functions seem to depend upon this type of connection (1, 4,

\*Received in the Editorial Office on November 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>Based on a dissertation submitted to the Graduate Division, University of California in partial fulfillment of the requirements for the Ph.D. degree. The author is greatly indebted to Professor Rheem F. Jarrett, under whose direction the research was conducted, and to Professor David Zeaman of the University of Connecticut who read the manuscript and made many helpful suggestions.

5, 6, 7, 10, 12, 21, 29, 30, 32, 36, 37, 46), while in others they do not appear to be operative (2, 8, 9, 10, 11, 14, 31). Related behavioral studies have not supported the notion of a simple transcortical system necessary for mammalian behavior. In several of Lashley's experiments rats with long cortical incisions have done very well in maze learning and retention (19, 25, 27), and Sperry (38, 39) has shown that separation of motor cortex in monkeys has little if any effect on movements.

In each of two closely related articles published in the early forties, Lashley presented data concerning both of the major points under consideration. The first paper (26) was interpreted as confirming, under carefully controlled conditions, that loss of the striate cortex has a much more serious effect on maze performance than loss of vision and consequently that striate cortex must be considered to have a nonvisual function important to maze learning. Yet, data in the second paper (27) showed that it was possible to account for the behavioral decrement in maze learning due to cortical lesions in terms of sensory loss alone without appeal to any additional cortical function. Lashley noted this, but felt that the results favoring the extra-sensory cortical function were clear cut. We shall try to show that analysis of the first experiment renders it inconclusive.

Some of the material from the earlier experiment appeared to require transcortical communication for maze learning, while the later one showed that such a system was at least not necessary. Lashley suggested that transcortical connections might constitute one of alternative sets of available pathways.

As a consequence of the conflicting nature of evidence in this field the present study was designed to obtain further information concerning the importance in complex learning of (a) the special "non-sensory function" of sensory cortex and (b) transcortical communication.

## B. EXPERIMENTAL DESIGN

Certain outcomes of an experiment using the following three groups would be sufficient to answer the question of the importance of transcortical connections in complex learning.

1. A group of normal rats trained to some criterion on a maze.
2. A group of rats with a given cortical area removed and similarly trained.
3. A group of rats with the same area isolated from all ventral connections and trained as above.

If all three groups performed the same then no conclusion could be reached



other than that the entire area was unimportant. If, on the other hand, Groups 2 and 3 differ, then the difference, regardless of which group it favors, must be due to the functioning of the partially isolated cortex in Group 3; the only way it could function effectively would be by means of transcortical connections of some type. For this particular question it would make no difference which part of the cortex was used so long as it was important in learning, and according to Lashley (19) all areas are important.

If the operations in Groups 2 and 3 were made in the visual area, then by addition of one more group we can obtain information concerning the non-sensory function of the visual area. This group would be:

4. A group of rats blinded by enucleation of the eyes and then trained to criterion.

If Group 2 learned significantly less well than Group 4, we would have further evidence for Lashley's claim of non-sensory function of the visual cortex. Equivalent performance of the two groups would not support such a theory.

As a control on the effectiveness of the undercutting operation in blinding the animals, a fifth group could be added. This would be:

5. A group of rats with the brains operated as in Group 3, with the eyes enucleated, and trained to criterion. If no difference were observed between Group 3 and Group 5 then these data could be pooled. However, if the two groups were not equal, their data would be difficult of interpretation.

The five groups used in the experiment will henceforth be designated as: Normal (N), Lesion (L), Undercut (U), Enucleate (E), and Undercut Enucleate (UE).

### C. METHOD

#### 1. *Subjects*

The subjects were 123 naïve, hooded, male rats of the Tryon M<sub>2</sub> strain obtained from the University of California Psychology Department animal colony. At the time of operation the age range was from 92 to 115 days. In so far as possible litter mate control was used and weight distributions of the groups were equated. Within these restrictions rats were randomly assigned to groups.

It was desired to have groups of approximately 20 animals each at the end of the experiment. To allow for losses, the operated groups were each assigned 26 animals, the enucleates 23, and the normals 22. Five animals with microphthalmia and two runts were discarded before the experiment be-

gan. Four died as a result of operations and three from excess anaesthesia. The records of seven others were not used, one because the rat was injured in the maze, two because the operations were unsatisfactory, and four because the animals died and their brains were not recovered. The 102 animals in the final groups were distributed as follows, Normal 21, Undercut 18, Lesion 22, Enucleate 21, and Undercut Enucleate 20.

## 2. Apparatus

The Tryon automatic maze was selected as the learning problem (40). The maze is a 17-unit alley *T*-maze with an "automatic" delivery table and automatic recording of errors. The running procedure has been adequately described by Tolman, Tryon, and Jeffries (40), and by Tryon (41, 42, 43).

## 3. Procedure

After the appropriate operations were performed, the animals were trained in the maze. A minimum of 14 and a maximum of 20 days elapsed between the operations and the beginning of training. Since there was an eight-day preliminary training period the time between the operations and the start of the learning was from 22 to 28 days. Enucleations in the Undercut Enucleate Group were carried out 7 to 10 days after the undercutting operation. All animals were allowed 15 to 16 days recovery from the enucleations to the start of the learning period.

The training in the Tryon maze was conducted in the standard manner with one trial run per day. Approximately 24 hours of combined hunger and thirst was used as motivation. At the end of the run the rats were given four teaspoonsful of wet mash, their total ration for the day. The average weight loss for the period in the maze was 27.7 grams. Animals were run for a minimum of 22 and a maximum of 30 learning trials. After the learning period in this maze they were rested and allowed to regain any lost weight. They were then run on three other experiments. Because of this, it was approximately nine months from the time of operations until the brains were recovered.

Two error scores were obtained. The first was the total errors for Days 2 to 22 inclusive (the standard measure for the Tryon maze). The second was the total errors made in reaching a criterion of performance. The choice of criterion was based upon the results of running other rats in this maze. It is impossible for most animals to reach a consistent errorless performance within a reasonable length of time; so a criterion of errorless performance is not practical. For the purpose of the second analysis (errors to criterion),

animals were run until they had made two errors or less on each of three successive days. Animals not reaching this criterion were run for a total of 30 days. Time scores were also obtained and total time scores for Days 2 to 22 will be presented.

#### 4. *Surgical Procedures*

Operations were performed under combined nembutal and ether anesthesia, enucleations under ether alone.

To meet the requirements of the experimental design of "undercutting" and removing the visual area of the cortex, the following operative techniques were used. The undercutting knife was a simple *L*-shaped instrument made of No. 9 piano wire, .022 inches in diameter. The small part of the *L* was 2.5 mm. long with the leading edge sharpened. A dental trephine 3.0 mm. in diameter was used to drill a hole in each side of the skull just anterior to the occipital suture and lateral to the sagittal suture. After the skull had been drilled a 2.5 mm. cut in the meninges, extending across the lateral diameter of the drilled hole, was made with a very fine scalpel. Then, by means of the Brown-Henry goniometer-manipulator (3) the knife was placed over the cut at the surface of the cortex with the blade as nearly normal to the surface as possible. The knife was then dropped 2 mm. into the cortex through the cut (the maximum depth of the visual cortex, 26), rotated 360 degrees and removed through the cut. This procedure was then repeated on the opposite side. The operation was designed to give two "undercut" circles of cortex 5 mm. in diameter which is approximately the length of the visual cortex (22, 24). Some damage was to be expected beyond the cut (27), and it was felt that this would help compensate for inaccuracies in drilling, individual differences in areas in the animals, etc. It was hoped that most of the visual area and some adjacent tissue would be undercut.

The "lesion" operations were performed with a *U*-shaped knife made of the same material as that used for the undercutting knife. The bottom and "moving" side were sharpened on one edge. The procedure was similar to that outlined above. Two holes were drilled with a larger trephine, 7.7 mm. in diameter. For most animals these holes extended from the center of the skull to the lateral edge of the dorsal surface. On a few rats with smaller skulls they actually overlapped in the center. They were placed as close to the sagittal suture as possible. A cut was made in the cortex from the medial side of the drill hole 2.5 mm. laterally. The knife was dropped 2 mm. into this incision, rotated 360 degrees and taken out. A cylinder of cortex approximately 5 mm. in diameter was then removed.

### 5. *Histology*

Immediately after each brain was extracted it was photographed.<sup>2</sup> The dorsal view and both lateral views were obtained, but only the dorsal view was used in the analysis. The photographs were projected on a three-inch scale to give reference points. All pictures were enlarged by the same factor, and lesions were plotted directly on the photographs. Brain sections of .020 mm. approximately  $\frac{1}{4}$  mm. apart were prepared and stained with thionin.

The size of the lesions and undercuttings was measured with the aid of a microscope especially adapted to an indicator gauge (17). With this instrument it was possible to determine the size and position of the lesion laterally to the nearest .001 inch. Measurements were transferred to the pictures by means of proportional dividers.

Levels of the lesions were identified by reference to the observable exterior damage and by a comparison of the levels of internal landmarks with those of several normal animals. Destruction of the cortex was set at the point where approximately either the upper or lower 50 per cent was very severely damaged. In most cases the beginning and ending of the destruction was quite clear. Damage in the fiber layers was identified by an abnormal number of small supporting cells in the fiber layer, a hole in the area, abnormality in the lower cortical cells, abnormality in the hippocampal cells, and distortion in the lower cortex or the hippocampus.

## D. RESULTS

### 1. *Anatomical Data*

*a. Lesion Group.* It can be seen in Figure 1 that the outlines of the lesions ranged roughly from circles to ellipses, the latter being more predominant. Since the operating instrument was designed to make circular lesions this requires explanation. It was noted during the operations that the meninges offered resistance to the cutting edge, causing distortion of the cortex as the knife pulled against them. This resistance appeared to make the lesions conform somewhat to the shape of the hemisphere; hence, the rough ellipses.

All of the lesions overlap other functional areas but the major damage is in the visual area. The principal extrastriate damage is anterior and medial to the striate area affecting the parietal, peristriate, and retrosplenial areas. Correspondingly the lateral and posterior portions of the striate area were most often unaffected by the lesions.

<sup>2</sup>Four rats were killed before the camera was ready, and so no photographs were obtained. These rats' lesions were plotted on photographs of normal brains.



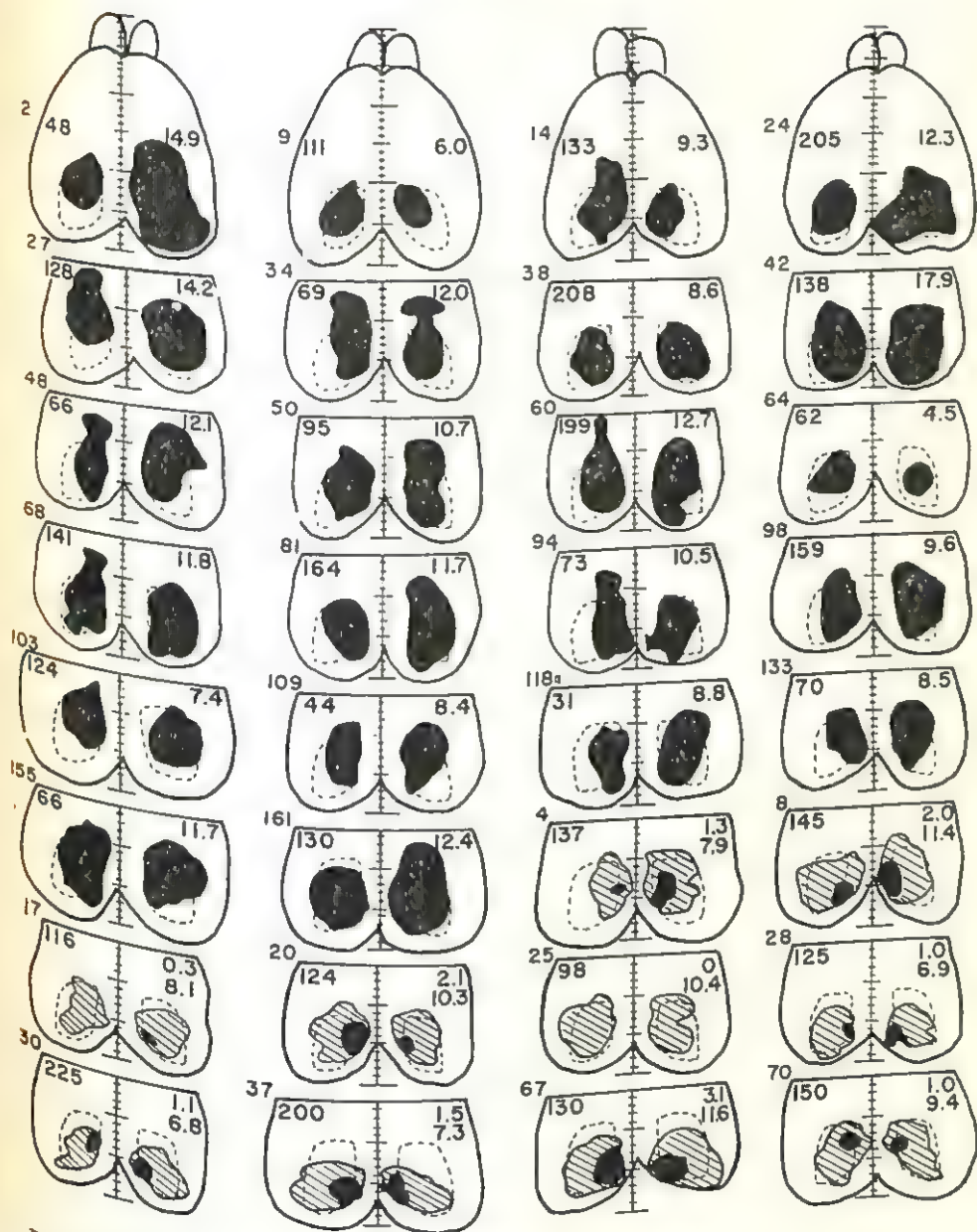


FIGURE 1  
 Diagrams of Lesion Animals 2 through 161 and Undercut Animals 4 through 70. Lesions are represented by solid black. Undercuttings are crosshatched. Visual areas are indicated by broken lines. Animal numbers are to the left of the diagrams. Total error scores appear within the diagrams on the left. Per cent of total cortex destroyed is at upper right with per cent undercut just below.

The average cortical destruction in this group was 10.7 per cent of the total cortex. The range was from 4.5 to 17.9 per cent.

*b. Undercut and Undercut Enucleate Groups.* The areas undercut in these two groups were somewhat smaller than the areas destroyed in the

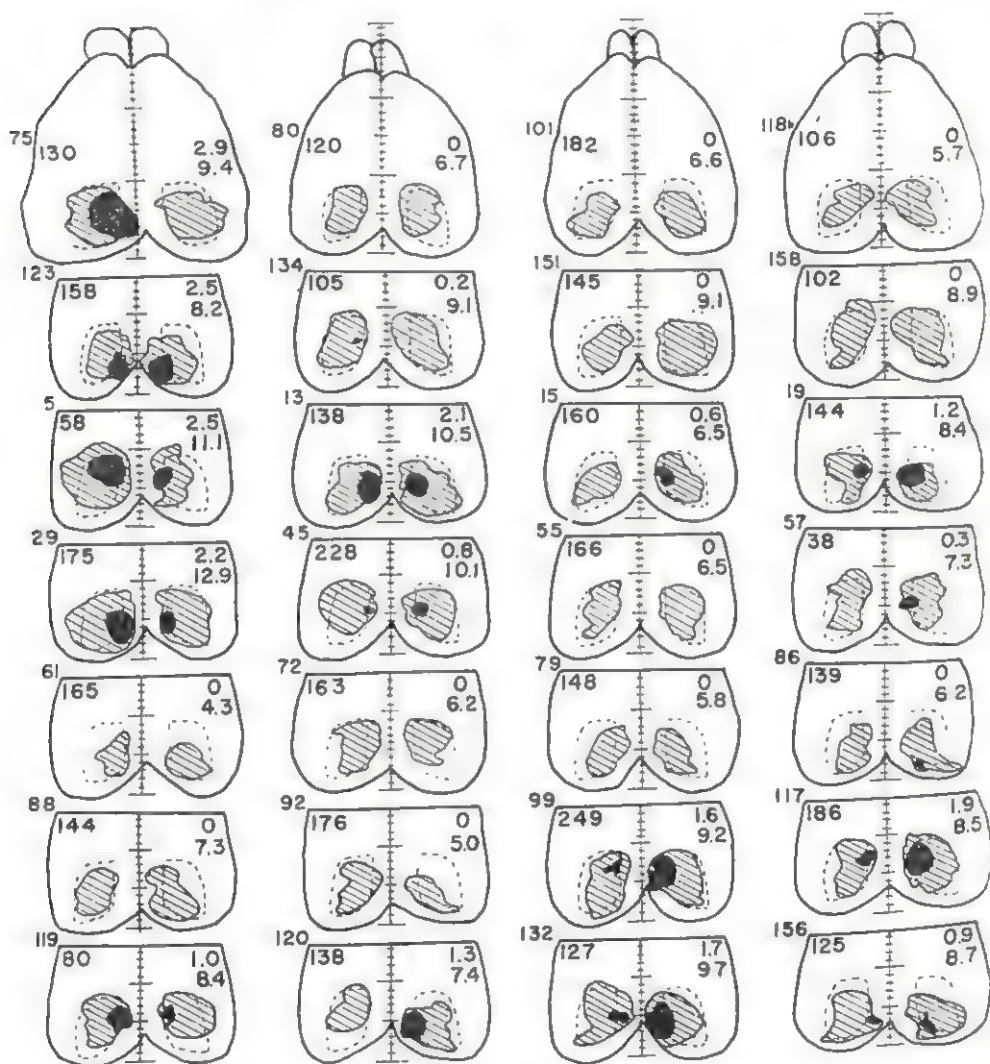


FIGURE 2

Diagrams of Undercut Animals 75 through 158 and Undercut Enucleate Animals 5 through 156. Lesions are represented by solid black. Undercuttings are crosshatched. Visual areas are indicated by broken lines. Animal numbers are to the left of the diagrams. Total error scores appear within the diagram on the left. Per cent of total cortex destroyed is at upper right with per cent undercut just below.

Lesion Group. The average for the Undercut Group was 8.5 per cent of the total cortex (range from 5.7 to 11.6 per cent), for the Undercut Enucleate Group it was 8.0 per cent (range: 4.3 to 12.9 per cent). The areas of actual destruction averaged 1.1 per cent (range: 0 to 3.1 per cent) in the undercuts, and .9 per cent (range: 0 to 2.5 per cent) in the undercut enucleates.

The undercuttings were shaped more like rough circles than rough ellipses, since the knife did not have to cut through the meninges. However, the cortex and underlying tracts were elastic enough to "give" to the blade, so that the shapes were not so regular as had been expected. The undercut areas cannot be considered as completely isolated as was originally planned. In many cases the blade went so deep as to leave a layer of normal fibers between the cut and the cortex. The amount of cortical destruction in these groups was insignificant and there was no overlap with the Lesion Group.

c. *Subcortical damage.* The principal subcortical structure damaged as a result of the operations was the hippocampus. Damage to each lobe was rated on a five-point scale (26). Two raters judged the damage independently and any discrepancies between the ratings were checked by reference to the slides. For the Lesion Group, eight animals had hippocampal damage on the right side. Of these one was rated 3, two rated 2, and five rated 1. On the left side in the same group, eight had damage, one rated 4, three rated 3, one rated 2, and three rated 1.

Similarly for the undercuts there were 10 with right hippocampal damage, two rated 3, one rated 2, and seven rated 1. On the left side there were 18, one rated 4, one rated 3, six rated 2, and ten rated 1.

In the undercut enucleates there were eight with damage on the right, one rated 4, one rated 3, two rated 2, and four rated 1. The left side had 19 cases, one rated 4, two rated 3, six rated 2, and ten rated 1.<sup>3</sup>

Damage to subcortical structures other than the hippocampus was largely limited to the colliculi and their tracts and generally was not extensive.

In the Lesion Group one animal had unilateral damage to both colliculi classified as slight to very slight.

Three animals in the Undercut Group had bilateral damage to the superior colliculi rated medium to slight. One had additional unilateral damage to the inferior colliculus. One undercut had unilateral damage to both

<sup>3</sup>Tables giving individual data on cortical, hippocampal, and tectum damage, and performance have been deposited with the American Documentation Institute. Order Document No. 4433 from American Documentation Institute, 1719 N. St. N. W., Washington 6, D. C., remitting \$1.25 for 35 mm. microfilm or \$1.25 for 6- by 8-in. photocopies.

colliculi, medium to slight, and one had severe damage to the left superior colliculus.

In the Undercut Enucleate Group 10 animals had some tectum damage. Two had bilateral superior collicular damage rated medium to very slight with slight involvement of one inferior colliculus. Five cases involved damage to both colliculi on one side. In four of these the damage was considered slight to very slight while in one it was medium to severe. Three animals had only one body damaged, two judged medium, and one very slight.

## 2. Performance

Three measures of performance were considered: total errors for Days 2 to 22, total errors for Days 2 to criterion, and total time scores for Days 2 to 22. From Table 1 it can be seen that enucleation by itself obviously

TABLE 1  
MEAN SCORES OF THE FIVE GROUPS

Group	<i>n</i>	Total errors Days 2-22	Total errors To criterion	Total time Days 2-22
Enucleate	21	111.3	124.2	2,551
Lesion	22	112.0	127.8	2,815
Normal	21	124.8	141.0	3,020
Undercut	18	138.8	173.7	3,091
Undercut Enucleate	20	147.4	186.5	3,250

caused no learning decrement in terms of any of the criteria. Nor are any of the differences between the undercuts and undercut enucleates statistically significant. Consequently the undercuts and the undercut enucleates will be treated as one group as originally proposed. This group will hereafter simply be called the Combined Undercut Group (*CU*).

Consider now the Days 2-22 scores. An analysis of variance of the four groups, *E*, *L*, *N*, and *CU*, yields an *F* of 3.03 (with 3 and 98 degrees of freedom) which has a probability between .01 and .05.

There are two comparisons we wish to make between our groups. The first is between the effect of peripheral blinding and "cortical" blinding. It is obvious that the difference between the Enucleate and the Lesion Groups is insignificant. Both groups are better than the Normals, although this is also not significant ( $F = 1.01$ ). The second comparison is that between the effect of removal of an area and "undercutting" of a similar area. This can be obtained by comparing the Lesion and Combined Undercut Groups. Analysis of variance gives  $F = 6.18$  (with 1 and 58 *d.f.*) again with a probability between .01 and .05.



Analyses based on total errors to criterion yield results comparable to the previous ones. The overall analysis of variance for the four groups with this criterion gives an  $F$  of 3.82 (for 3 and 98 *d.f.*) with  $p$  less than .05 but greater than .01. Again the Enucleate and Lesion Groups do not differ. They are both better than the normals but not significantly ( $F = .53$ ). The comparison of the Lesion and Combined Undercut Groups resulted in an  $F$  of 7.43. The probability of this  $F$  (for 1 and 58 *d.f.*) is less than .01.

The trend of the means for total time is the same as for the other criteria, but none of the differences are significant.

### 3. *Effects of Subcortical Damage on Performance*

The results of the Lesion Group in this experiment and Lashley's Group III in his 1943 experiment (26) give consistent evidence concerning the effect of damage to the hippocampus upon maze performance. Both of these groups had such damage and performed as well as, or better than, normal animals. It appears reasonable then to disregard the effect of hippocampal damage to the Undercut and Undercut Enucleate Groups. But, in so doing we should note that the hippocampal damage in the Combined Undercut Group was more severe than in the Lesion Group, and that while the damage in Lashley's Group III was quite severe the number of animals involved was small.

Analysis of the damage to the corpora quadrigemina is more difficult. Within our experimental groups we shall compare animals having such damage with those lacking it. The comparison will be given in terms of the total errors for Days 2 through 22. Only two of the animals in the Lesion Group had such damage, and they both performed better than the mean of the rest of their group. Since the mean of this group is superior to the normals they are both superior to the average normal animal.

The animals in the Combined Undercut Group with damage to the colliculi averaged 156.9 errors while those without such damage averaged 134.4 errors. A significance test gives a value of  $t$  having a probability greater than 10 per cent. The individual scores yield some evidence to support the argument that the difference is a chance difference. Every group has one or two scores over 200, so such scores must be considered in the normal range. The Undercut Group had two such scores, but these were two of the five rats in this group which had damage to the colliculi, so they had a considerable effect in raising the mean of their subgroup. It is also worth noting that one rat who had severe subcortical damage to all four of the

colliculi had a score of 132, considerably below the mean of the Combined Undercut Group.

The evidence suggests that the damage to the colliculi may not have affected the learning of rats in this maze, but cannot be considered conclusive.

## E. DISCUSSION

### 1. *Non-Sensory Function*

The first question to be considered is that of a nonvisual function of the visual cortex in maze learning. The results of the Normal, Enucleate, and Lesion Groups demonstrate no non-sensory function in the learning of this maze. The centrally blinded group performed exactly as the peripherally blinded group, and both of these performed slightly better than the normals.

A brief review of previous work on the problem will aid us in evaluating these results. As indicated earlier both Lashley and Tsang obtained consistent results suggesting that loss of a specific sense by destruction of the sense organ had a much less serious effect upon maze performance than loss of the same sense through destruction of its cortical projection area. Lashley used these results to support the argument that the sensory projection areas must have in addition to their sensory function a non-sensory function important to maze behavior. This additional function he described as "dynamic" or "facilitative."

Finley's criticism (13) of these early experiments was aimed chiefly against the use, in critical experiments, of data based upon very small numbers of animals, and the fact that lesions considered as effectively falling in a given sensory area overlapped other areas as well and often were accompanied by subcortical damage. She says: "The destruction was, in no one of the experiments under discussion, limited to the visual projection area of the cortex, and the results do not then support the conclusion that destruction of a sensory projection area of the cortex produces a disturbance of function from nine to sixty times as great as destruction of the corresponding peripheral mechanisms" (13, p. 207). Finley attacked the problem with a new experimental design. She had two groups of animals, one described as having lesions restricted to the visual area and a second with lesions of "approximately" the same size but involving more than one functional area. The former were found to do no worse in post-operative runs on a maze than they had previously done in the dark; hence, central deprivation had no greater effect than the peripheral elimination of visual cues. The second group however did show a post-operative deterioration in performance, presumably due to the involvement of nonvisual structures.

Lashley (26) accepted Finley's criticism that lesions in his earlier studies had gone beyond the visual area, but pointed out that only subsequent to these studies had the visual cortex been adequately plotted (22, 24). He disposed of Finley's experiment by pointing out that the lesions of her first group were not in fact limited to the visual area and, moreover, were so small that no observable effects should have been expected. Her second group he describes as having lesions which when compared to those of the first group are found to invade no greater nonvisual areas of the cortex but are approximately 50 per cent larger (presumably in the visual area) and thus caused deterioration in performance.

The experiment he then presents will be considered in some detail. Four experimental groups were used. Group I (normal animals) was trained on the Lashley maze; the animals' eyes were enucleated and the group was retrained. Group II animals were blinded by enucleation and then trained on the maze. The animals of Groups I and II were then combined, operated upon in the visual area and retrained. Group III was made up of normal animals trained, centrally blinded by cutting the posterior thalamic radiations with minimal damage to the overlying cortex, and retrained.

If we consider the results of Group III first, we find that the retention of the maze habit after operation equals that of Group I. The combined group whose animals had the visual cortex destroyed with "invasion of adjacent structures similar to that involved in section of the radiation," showed "ten" times as great deterioration in maze retention as Group I. These data together seem to support quite clearly the argument for a non-sensory function of the visual cortex.

In the same experiment, Lashley makes clear, however, that any conclusions of this kind are theoretically unpredicted. He states: "Ideally a conclusive test for nonvisual functions of the visual cortex should consist of training peripherally blinded animals, then measuring the effect of lesions strictly limited to the area striata. Such an experiment is practically impossible since the *minimal size of lesion which produces measurable deterioration of maze performance approximates the area of the visual cortex . . .*" (italics mine) (26, p. 434). Yet, he concludes that total destruction of the visual cortex, combined with other damage, produces severe deterioration in maze performance, at least 10 times (revised to 8 in 1944, 27) as great as that following peripheral blinding, and that analysis shows that the damage to structures other than the striate area did not contribute significantly to the loss of the maze habit.

Thus we have the removal of the visual cortex, alone responsible for,

severe deterioration in maze performance although this area is so small that it is questionable whether it could produce measurable deterioration. Here is a dilemma. Either his assertion as to the minimal size of lesions which will cause deteriorated performance must be wrong, or his reasoning leading to the conclusion that damage to adjacent areas was unimportant must be faulty.

There is considerable evidence that the performance decrement due to lesions the size of the striate area would be quite small (18, 19, 28, 33, 34).<sup>4</sup>

Let us, then, examine the reasoning leading to the conclusion that damage to areas other than the visual had no effect upon performance. The first argument asserts this damage was "similar" in the Combined Group and Group III. Since Group III showed no deterioration in performance it is concluded that such damage could not contribute to the deteriorated performance of the Combined Group. It may be true that the subcortical damage in the two groups was "similar," but it is clear from the diagrams of cortical damage Lashley presents that there is considerably more *extrastriate* cortical damage in the Combined Group than in Group III and further that the damage is not in comparable areas in the two groups. Consequently this argument cannot be considered satisfactory.

A second argument concerning the inconsequential rôle of the nonvisual areas proceeded in two steps. First Lashley presents evidence that any one of a number of specific extrastriate areas had no measurable effect on maze behavior. This is followed by evidence that the *cumulative* effects of damage to a number of nonvisual areas had no effect on behavior. Here the evidence is weak and totally inconsistent with his own previous results (19, see especially Table XVI p. 106) and his theory of mass action.

Lashley has presented us, here, with an experiment in which the average cortical damage to animals was 20.0 per cent. Previous tests of the same performance in the same maze (19) have shown that lesions of this size in various cortical areas could be expected to result in severely deteriorated performance, while smaller lesions (10 per cent or less) have no appreciable effect. Nonetheless, Lashley has tried to show that the severe deterioration found is to be attributed to approximately half the damage while the other

---

<sup>4</sup>Lashley has outlined the visual area by a series of lesions which gave no degeneration in the lateral geniculate pars dorsalis. There is a slight discrepancy between the areas so outlined (22, 24) and the area he presents on his diagrams as the visual area. Since the discrepancy is unaccounted for we have used the area represented by the original data and have calculated it to be 8.0 per cent of the total cortex.



half is inconsequential. Why the mass action effect of the visual area is so crucial while that of the other areas is insignificant is not explicated.

A conclusion supported by all of these criticisms states that the greater deterioration in performance of Lashley's cortically blinded group can be ascribed to the cumulative effect of striate and extrastriate cortical damage without invoking a non-sensory "facilitative" function of the visual cortex. It may be true that the striate area, as well as other cortical areas, has important integrating, or facilitative or dynamic functions, but it has not been demonstrated by Lashley's experiment nor could it be expected to be demonstrated by any experiment of this design. Our experiment has shown that, given lesions which approximate the size of the visual area (average, 10.7 per cent as contrasted to Lashley's average, 20.0 per cent) and lie principally in the visual area, no deterioration in performance can be found in this complex learning problem.

## 2. *Transcortical Connections*

The second question to be considered is the importance of transcortical connections in complex maze learning. Relevant evidence is given by the animals in Lashley's Group III. These animals, in whom the thalamic radiations to the visual area were cut, performed as well as enucleated rats. Lashley suggested that this indicated the visual area was functioning, and, further, functioning in a horizontal direction. This would support the idea of some system of transcortical connections.

It is clear from our discussion of the visual area that even if it were not functioning no difference between the groups should be expected. Thus, these results give us no evidence either for or against a system of transcortical connections.

In 1944 Lashley (27) directly attacked the problem of transcortical associative processes in learning and retention. He did this by making long incisions in the cortex of rats with as little damage to the cortex or other structures as was possible. All rats had symmetrically placed incisions one in each hemisphere. An attempt was made to separate major functional areas by the incisions, but later analysis indicated that in general they went through the major areas rather than between them. The average length of the incisions was 50 per cent of the length of the hemisphere. If intracortical associations were mediated by fibers running horizontally through the cortex, Lashley felt that these incisions should seriously disrupt such processes as complex maze learning or retention. Although the performance of the operated animals was definitely inferior to that of normals, he found that it was

possible to account for any deterioration in performance either on the basis of the cortical damage caused by the incisions or the sensory defect due to degeneration in the thalamus caused by the incisions cutting thalamic radiations. This clearly gave no support to a system of transcortical connections. However, it is not conclusive evidence against such a system if one accepts the possibility of multiple pathways, not all of which are likely to be severed by a straight cut.

In our experiment we found a difference between the Lesion Group and the Combined Undercut Group significant at the 5 per cent level of confidence. If we accept the analysis of subcortical damage as showing that it did not contribute to the behavioral difference, we can look for the source of the difference in other characteristics of the two groups. The basic question concerns the effect of the undercutting operation. Had the undercutting rendered the overlying cortex nonfunctional neither group would be expected to be measurably affected. If there were a measurable effect, the lesions should show the poorer performance since they have the larger involved areas.

Undercutting could interfere with fiber tracts to areas other than those directly above the cut thus hindering the function of larger areas of cortex. While such damage unquestionably occurred, the damage to fiber layers directly subjacent to the cortex is greater in the lesions than in the undercuts; so, the results should have been reversed.

A tentative hypothesis consistent with the data would be that the undercutting caused malfunctioning of its overlying cortex and that this interfered with the functioning of other areas. The principal difference between the two groups is that the undercuts have an area of cortex which the lesions lack. The performance of the group with the cortex is poorer. This suggests that the cortex is exerting a negative effect upon learning. Now even when the undercut cortex was not completely isolated from below there was interference with its normal projections. It is possible that functional cortical tissue lacking its normal innervations may malfunction (16) and so interfere with other areas in a complex learning problem. There is evidence that undercut cortex may malfunction (11) and of course abundant evidence that at least some types of abnormally functioning cortex affect behavior (35).

This hypothesis, which requires a means of transmission of the abnormal activity, is consistent with a system of transcortical connections. Since the isolation of undercut areas was not complete it can be considered only suggestive.

## F. SUMMARY AND CONCLUSIONS

Results of the experiments reported above show that for learning in the complex alley maze employed:

1. Neither peripheral blinding by enucleation of the eyes nor destruction of most of the visual projection area of the cortex of the rat, together with slight damage to other cortical and subcortical areas, produces any deterioration in learning performance and may result in slightly improved performance.
2. No evidence was found to support the hypothesis that the visual cortex has any function in learning other than sensory projection.
3. Cutting the fiber layer subjacent to the visual cortex, with minimal damage to the cortex and some damage to the hippocampus and colliculi, results in greater deterioration in maze learning performance than does removing similar cortical areas, with less hippocampal and colliculi damage.
4. The deterioration in performance of the undercut animals can be explained in terms of an abnormal functioning of the undercut areas which interferes with other cortical areas.
5. Such an explanation would be consistent with Lashley's suggestion that transcortical connections are important in the normal operation of a cortical area in learning.

## REFERENCES

1. ADRIAN, E. D. The spread of activity in the cerebral cortex. *J. Physiol.*, 1936, **88**, 127-161.
2. BARTLEY, S. H., & BISHOP, G. H. Factors determining the form of the electrical response for the optic cortex of the rabbit. *Amer. J. Physiol.*, 1933, **103**, 173-184.
3. BROWN, C. W., & HENRY, F. M. The central nervous mechanism for emotional responses: III. A combination headholder and goniometer-manipulator for controlling movements of a point electrode within the brain. *J. Comp. Psychol.*, 1935, **20**, 53-58.
4. BUCY, P. C. Electrical excitability and cytoarchitecture of the premotor cortex in monkeys. *Arch. Neurol. & Psychiat.*, 1933, **30**, 1205-1225.
5. ———. The relation of the premotor cortex to motor activity. *J. Nerv. & Ment. Dis.*, 1934, **79**, 621-630.
6. BUCY, P. C., & FULTON, J. F. Ipsilateral representation in the motor and premotor cortex of monkeys. *Brain*, 1933, **56**, 318-342.
7. DUSSER DE BARENNE, J. G., GAROL, H. W., & McCULLOCH, W. S. Functional organization of sensory and adjacent cortex of the monkey. *J. Neurophysiol.*, 1941, **4**, 324-330.
8. ———. Physiological Neuronography of the cortico-striatal connections. *Res. Publ. Ass. Nerv. & Ment. Dis.*, 1942, **21**, 246-266.
9. DUSSER DE BARENNE, J. G., & McCULLOCH, W. S. Functional organization in the sensory cortex of the monkey (*Macaca Mullata*). *J. Neurophysiol.*, 1938, **1**, 69-85.

10. ———. Functional interdependence of sensory cortex and thalamus. *J. Neurophysiol.*, 1941, 4, 304-310.
11. ———. Suppression of motor response obtained from area 4 by stimulation of area 4s. *J. Neurophysiol.*, 1941, 4, 311-323.
12. DUSSEY DE BARENNE, J. G., & MARSHALL, C. S. On a release phenomenon in electrical stimulation of the "motor" cerebral cortex. *Science*, 1931, 73, 213-214.
13. FINLEY, C. B. Equivalent losses in accuracy of response after central and after peripheral sense deprivation. *J. Comp. Neurol.*, 1941, 74, 203-237.
14. GERARD, R. W. The interaction of neurons. *Ohio J. Sci.*, 1941, 41, 160-172.
15. HONZIK, C. H. The sensory basis of maze learning in rats. *Comp. Psychol. Monog.*, 1936, No. 13.
16. JACKSON, J. H. Evolution and dissolution of the nervous system. Croonian Lectures. *Brit. Med. J.*, 1884, 1, 660-663.
17. JARRETT, R. F. An experimental study of the inter-relations between lesion size, motivation and problem difficulty in maze-learning by the rat. Unpublished doctor's dissertation, University of California, 1947.
18. LASHLEY, K. S. Studies of cerebral function in learning: VII. The relation between cerebral mass, learning, and retention. *J. Comp. Neurol.*, 1926, 41, 1-58.
19. ———. Brain Mechanisms and Intelligence. Chicago: Univ. Chicago Press, 1929.
20. ———. Cerebral control versus reflexology: A reply to Professor Hunter. *J. Gen. Psychol.*, 1931, 5, 3-20.
21. ———. Mass action in cerebral function. *Science*, 1931, 73, 245-254.
22. ———. The mechanism of vision: VIII. The projection of the retina upon the cerebral cortex of the rat. *J. Comp. Neurol.*, 1934, 60, 57-79.
23. ———. The mechanism of vision: XII. Nervous structures concerned in the acquisition and retention of habits based on reactions to light. *Comp. Psychol. Monog.*, 1935, 11, 43-79.
24. ———. Thalamico-cortical connections in the rat's brain. *J. Comp. Neurol.*, 1941, 75, 67-121.
25. ———. The mechanisms of vision: XVII. Autonomy of the visual cortex. *J. Genet. Psychol.*, 1942, 60, 197-221.
26. ———. Studies of cerebral function in learning: XII. Loss of the maze habit after occipital lesions in blind rats. *J. Comp. Neurol.*, 1943, 79, 431-462.
27. ———. Studies of cerebral function in learning: XIII. Apparent absence of transcortical association in maze learning. *J. Comp. Neurol.*, 1944, 80, 257-281.
28. LASHLEY, K. S., & WILEY, L. E. Studies of cerebral function in learning: IX. Mass action in relation to the number of elements in the problem to be learned. *J. Comp. Neurol.*, 1933, 57, 3-55.
29. LEAO, A. A. P. Spreading depression of activity in the cerebral cortex. *J. Neurophysiol.*, 1944, 7, 359-390.
30. LEAO, A. A. P., & MORISON, R. S. Propagation of spreading cortical depression. *J. Neurophysiol.*, 1945, 8, 33-45.
31. LIBET, B., & GERARD, R. W. Steady potential fields and neuron activity. *J. Neurophysiol.*, 1941, 4, 438-455.
32. MCCULLOCH, W. S. Modes of functional organization of the cerebral cortex. *Fed. Proc.*, 1947, 6, 448-451.
33. MAIER, N. R. F. Cortical destruction of the posterior part of the brain and its effect on reasoning in rats. *J. Comp. Neurol.*, 1932, 56, 179-214.



34. ———. The effects of cerebral destruction on reasoning and learning in rats. *J. Comp. Neurol.*, 1932, **54**, 45-75.
35. PENFIELD, W. G., & ERICKSON, T. C. *Epilepsy and Cerebral Localization*. Springfield, Ill.: Thomas, 1941.
36. ROSENBLUETH, A., BOND, D., & CANNON, W. B. The control of clonic responses of the cerebral cortex. *Amer. J. Physiol.*, 1943, **137**, 681-694.
37. ROSENBLUETH, A., & CANNON, W. B. Cortical responses to electric stimulation. *Amer. J. Physiol.*, 1942, **135**, 690-741.
38. SPERRY, R. W. Horizontal intracortical organization in cerebral control of limb movement. *Proc. Soc. Exp. Biol. & Med.*, 1945, **60**, 78-79.
39. ———. Cerebral regulation of motor coordination in monkeys following multiple transection of sensorimotor cortex. *J. Neurophysiol.*, 1947, **10**, 275-294.
40. TOLMAN, E. C., TRYON, R. C., & JEFFRIES, L. A. A self-recording maze with an automatic delivery table. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 99-112.
41. TRYON, R. C. Individual differences at successive stages of maze learning. Unpublished doctor's dissertation, Univ. California, 1928.
42. ———. Genetics of learning ability in rats. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 71-89.
43. ———. Studies in individual differences in maze ability: I. The measurement of the reliability of individual differences. *J. Comp. Psychol.*, 1930, **11**, 145-170.
44. TSANG, Y.-C. The functions of the visual area of the cerebral cortex of the rat in the learning and retention of the maze: I. *Comp. Psychol. Monog.*, 1934, **10**, No. 4.
45. ———. The functions of the visual areas of the cerebral cortex of the rat in the learning and retention of the maze: II. *Comp. Psychol. Monog.*, 1936, **12**, No. 2.
46. WOOLSEY, C. N. Patterns of sensory representation in the cerebral cortex. *Fed. Proc.*, 1947, **6**, 437-447.

Department of Psychology  
University of Connecticut  
Storrs, Connecticut

## SOME "FIRST" SOURCES OF SEX INFORMATION AS REPORTED BY SIXTY-SEVEN COLLEGE WOMEN\*<sup>1</sup>

*College of Education, University of Oklahoma*

HENRY ANGELINO AND EDMUND V. MECH<sup>2</sup>

### A. THE PROBLEM AREA

Eventually, children of school age become extremely interested and somewhat curious over matters dealing with sex and quite naturally desire to gain further information on the subject. The first or initial source of information on matters of sex might well create attitudes in young boys and girls which could be quite permanent. Adults, of course, would hope that the youngsters receive accurate information concerning sex and, in addition, would concur that healthy attitudes toward the topic should be instilled in children.

However, Ramsey (3) questioned a sample of pre-adolescent and early adolescent boys regarding their initial source of sex information on such topics as Ejaculation, Nocturnal Emissions, Contraceptives, Menstruation, Masturbation, Intercourse, Prostitution, and Venereal Disease. It was found that a greater portion of "information" was obtained from other boys in their own age group than was furnished by either the mother or father of each boy. Horrocks (2) points out some undesirable consequences of the foregoing results by stating:

The result is often a hodge-podge of misinformation, half-true guesses or shrewd surmises, folklore, and out-and-out misconceptions which can only be cleared up by actual experience or a proper course of sex education. As a matter of fact, even actual experience fails to answer questions or clear up misconceptions, if we are to judge by the questions asked by adolescents of various ages and degrees of experience (p. 174).

In view of the increasing interest in sex education as it relates to the academic curricula, the present paper was concerned with the following two facets of the problem: (a) what "first" sources of sex information on a variety

\*Received in the Editorial Office on November 20, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>Portions of this paper were presented at the 1954 meeting of the Oklahoma Academy of Science held in Tulsa.

<sup>2</sup>The authors wish to express their gratitude to the Faculty Research Committee of the University of Oklahoma for financial support that made publication of this report possible.

of "topics" are listed by females and (b) what suggestions are given by these females for the improvement of sex education, either at home or in the school.

### B. PROCEDURE

Sixty-seven females who were enrolled in a "Human Growth and Development" course at the University of Oklahoma were used in the study. Each *S* was required to state to the best of her ability the initial source of "information," either accurate or inaccurate, on the following sex topics: Contraceptives, Venereal Disease, Prostitution, Masturbation, Intercourse, and Menstruation. As a further requirement *S* was asked to indicate what could be done, on the part of parents or the schools, to improve sex education instruction. It was felt that in this way each individual *S* might possibly divulge some particular deficiency in sex education as it pertained to her personally. All *Ss* were allowed one hour to complete their reports, which were turned in unsigned by the respondents.

### C. RESULTS

The important numerical data are summarized in Table 1. Table 1

TABLE 1  
FREQUENCY OF OCCURRENCE OF "FIRST" SOURCES OF SEX INFORMATION IN EACH OF SIX AREAS AS REPORTED BY 67 FEMALES

First source	Contraceptives	Venereal Disease	Prostitution	Masturbation	Intercourse	Menstruation
Female companions	31	10	24	11	19	9
Mother	8	12	9	4	20	48
Father	1	1	0	0	2	0
Both parents	0	0	3	0	0	0
Printed matter	8	21	22	27	13	5
Male companions	1	0	1	4	3	0
Other relatives (brother, sister, aunt)	6	3	1	0	3	3
School courses	5	7	4	10	3	2
Other sources and undecided	7	12	3	11	3	0
Total	67	66	67	67	66	67

shows the frequency of occurrence (*not* in percentages) of "first" sources of sex information for each of six topics as stated by 67 college women. No really meaningful comparison can be made at this time with the Ramsey study. One primary reason for this lack of comparability is that the *Ss* in

our study were required to "reflect back" over a much longer period of time than was the case in the Ramsey situation. Ramsey studied "pre-adolescent" boys while the majority of Ss in our sample were college sophomores approximately 19 years old. Data are now being gathered which will permit comparison with the Ramsey study.

However, the data in Table 1 in a broad sense follow the trend in the Ramsey study. Certainly, the effects of parents and the school are not very heavy, at least with respect to acting as an "initial" source of sex information. It would appear that with the exception of menstruation, girls received most of their sex information from their companions and from "printed matter." Interestingly enough, the father himself played almost no rôle at all in providing information.

As examples of the need for "sex" education consider the following excerpts taken from the verbatim reports of Ss used in our college sample:

"I think that the parents should have the biggest responsibility in informing their children of sex, but many parents do not perform this responsibility.

"My knowledge of sex was increased as I began to read articles and discuss it with the girls of my own age. Many times we were misinformed."

"As nearly as I can remember I had no sex education at all in elementary school. I would suggest that a program of this type on the grade school level be worked into the health course that is offered in many smaller schools."

"I think that not enough parents realize the importance of giving their children a good education on matters concerning sex. A few high schools have sex education classes, but I think this should be a required curriculum in all high schools. There are good sex education classes in college, but very few people take advantage of this opportunity to learn more and get more accurate information!"

"Most of the classes that 'frankly' discussed these subjects were in college. I think this type of education should be inserted into the school curriculum in junior high—and I think it should be required. In this way every child will have accurate information. If the course is treated as matter-of-fact, the old 'hush-hush' situation, as well as the other extreme, will be eliminated and sex can take its natural place in the educational system."

"I had many misconceptions. I don't think anyone ever gave me the wrong ideas, but I was pretty vague on what they didn't tell me. I think there should be more information given to the public. Most people won't buy these books, but if they were put out through the school, it would help. Parents should also be taught how to explain these things to their children."



"When I was wanting information about sex, I read numerous articles, but they all seemed to deal in generalities. I would have liked to know the facts in a way that I could understand them rather than hearing someone 'beat around the bush.' I would have liked to have the facts which I got in a freshman psychology class in college when I was a freshman in high school. I also think adult education classes would help parents in dealing with subjects of sex with their children."

"I think that parents should all take a course in how to educate their children about sexual matters. Too many parents leave their children too much in the dark concerning the necessary facts about sex. I feel as though the information we receive from other members of the "crowd" might well be too misleading. The parents should start telling their children a few facts of life before they are even in Jr. High. And when the child asks questions concerning sexual matters the parents should give them the truth in the language in which the child can understand."

The tenor of the foregoing remarks made by college women is in keeping with a recent statement by Blair, *et al.* (1) as follows: "There is much evidence that our young people are eager for instruction in such matters, [sex education]<sup>3</sup> but that schools have done very little in the way of giving effective guidance in this area" (p. 85).

The data are certainly clear in pointing out that young persons are not, in effect, receiving adequate and practical information, either at home or in the public school, on the variety of sex topics they are so interested in learning more about. It can be concluded then, with some margin of safety, that the needs of both boys and girls relative to proper sex information are not being adequately met either by the home or the school.

#### REFERENCES

1. BLAIR, G. M., JONES, R. S., & SIMPSON, R. H. *Educational Psychology*. New York: Macmillan, 1954. Pp. 601.
2. HORROCKS, J. E. *The Psychology of Adolescence*. Boston: Houghton Mifflin, 1951. Pp. 614.
3. RAMSEY, G. V. The sex information of younger boys. *Amer. J. Orthopsychiat.*, 1943, **13**, 347-352.

*College of Education*  
*University of Oklahoma*  
*Norman, Oklahoma*

<sup>3</sup>Brackets ours.

## THE AUTHORITARIAN PERSONALITY AS A STEREOTYPE\*

*I.A. Hospital, Brockton, Massachusetts*

BERNARD A. STOTSKY<sup>1</sup>

### A. THE PROBLEM

The authors of *The Authoritarian Personality* have explicitly stated that there is no direct relationship between authoritarianism and mental health (1, 2). Masling (6) cites four independent studies to show no relation between the two. However, in psychology courses, when we discuss the authoritarian individual, we freely quote the findings of studies (1, 2, 7) which characterize him as superstitious, stereotyped, cynical, destructive, sadomasochistic, ambivalent toward authority, punitive, full of aggression towards the weak, and maladaptive. By describing the authoritarian individual in such terms, one is led to wonder what conclusions students will draw concerning the relationship between authoritarian attitudes and poor mental health.

### B. PROCEDURE

In this study the author attempted to investigate among undergraduate psychology students the relationship which they perceived between the two. Three classes of undergraduate students in the Psychology of Adjustment course at Wayne University served as subjects. Of the 137 participating students, 78 were male and 59 female. Two scales, the Taylor Manifest Anxiety Scale (8) and the California *F* (ascism) Scale, were administered during two consecutive class periods, with the following instructions:

A research project is being conducted at Wayne University in an attempt to find out about some of the attitudes, feelings, and behavior characteristics of college students. We would very much like your cooperation in the project. Results are being used in research purposes only and your answers will be considered as confidences and respected as such. There are no right or wrong answers. It is merely necessary to indicate how you feel. It is very important that you answer frankly and objectively. The results will mean relatively little unless you express your opinions honestly.

\*Received in the Editorial Office on November 23, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.  
<sup>1</sup>The author wishes to express his appreciation to Professor Henry Weinberg of Boston University for his constructive suggestions concerning the paper.

The Taylor scale was administered first and then followed by three administrations of the *F* scale: (a) *Experimental Condition I*: *F* scale administered with instructions to answer it according to "your personal opinion, how you feel about these items." *Experimental Condition II*: *F* scale administered with instructions to answer it as "the neurotic individual, a person in poor mental health who showed anxiety symptoms, would answer it." *Experimental Condition III*: *F* scale administered with directions to answer it as "the normal person would."

All students in the classes participated. No sex differences were found for Anxiety scores or *F* scores. There were no differences between classes in response to the scales.

### C. FINDINGS

Is the belief that authoritarianism and neurosis are related widely held by the students in this sample? Comparison of scores between experimental conditions I and II, I and III, II and III shows differences significant beyond the .001 level. The students perceived themselves as manifesting

TABLE 1  
COMPARISON OF *F* SCALE SCORES OF 137 COLLEGE STUDENTS FOR THE THREE  
EXPERIMENTAL CONDITIONS\*

	Mean Diff.	CR	<i>p</i>
Exp. Cond. I and II (Self vs. Neurotic)	74.1	19.9	<.001
Exp. Cond. I and III (Self vs. Normal)	39.0	11.3	<.001
Exp. Cond. II and III (Nor. vs. Neur.)	35.1	9.1	<.001

\*Mean, Cond. I, 84.3.

Mean, Cond. II, 158.4.

Mean, Cond. III, 123.3.

far less authoritarian attitudes than either normals or neurotics. Neurotics were regarded by them as much more authoritarian individuals than normals. For 106 out of 137 students *F* scale scores attributed to neurotics were greater than those attributed to normals. In comparison with the students' own responses to the *F* scale neurotics were seen as giving higher *F* scores in 125 out of 137 cases. This indicates that there is a widespread belief among these students that neurotics are more authoritarian than either normals or themselves. Comparison of students' ratings for neurotics with ratings for normals shows a perceived relationship of 1.64 between high *F* scores and neurosis (biserial *r*).

### D. DISCUSSION

Why do students relate authoritarianism and poor mental health so strongly, especially when the facts appear to indicate otherwise? To un-

derstand why they do, the author feels that it is necessary to study this belief as a stereotype. By stereotype is meant a widespread belief which conforms very little to the facts it pretends to represent (3, 4, 5). It results from our defining first and observing second (4). In this study we have found there is a widespread belief that authoritarianism and poor mental health are highly interrelated. This belief persists in spite of repeated findings, including those by the authors of *The Authoritarian Personality*, that no such relationship exists. Thus, since the belief is widespread among these students and since it does not conform to the facts, it may be correctly classified as a stereotype.

It may be argued that the finding of no relationship in earlier studies does not necessarily mean that such a relationship does not exist in the present study among these students. To investigate this possibility, 33 students with the highest Taylor scores were selected and compared with the remaining 104 students. All 33 students showed anxiety scores of 21 or greater so that they could be regarded as moderately anxious. This group gave significantly lower *F* scores than the group made up of less anxious students in Condition I (self rating). This finding is in direct contradiction to the students' perception of the response of such individuals to the *F* scale (Condition II). The difference between the mean expected score for neurotics (158.4) and the score obtained by a group of anxious students (72.4) is significant beyond the .001 level. Moreover, no significant relationship was found between anxiety scores and *F* scores in any of the experimental conditions. The correlation between anxiety scores and *F* scores in Condition I (self rating) was  $-.092$  which is not significant. Comparison of individuals with low *F* scores with individuals with high *F* scores for anxiety scores showed no significant difference.

The 33 anxious individuals were no different from the rest in their perception of the neurotic's response to the *F* scale. The average *F* score for these individuals in Condition II was 156.3, for the remaining 104, 159.1. The difference was not significant. Thus, for the subjects of this study no relationship was found between authoritarian attitudes and poor mental health, the criterion for the latter being a high Taylor anxiety score. Our findings support those of previous studies concerning the relationship between the two.

What is responsible for this distortion in student attitudes? The manner in which students emphatically rejected authoritarian attitudes with respect to themselves and then proceeded to project them onto neurotics suggests, as Masling has so felicitously phrased it, a tendency to use the term



"authoritarian" as a mild profanity which one could use to describe other people, but never oneself. Why the uncritical acceptance of such a belief in spite of experimental evidence to the contrary? Can it be that we psychologists are responsible for fostering such a belief among our students? Have we by insinuation, speculation, and indiscriminating use of value terms influenced students to think in this manner? Are we unwittingly assisting them to develop a stereotype concerning the authoritarian while encouraging them to discard stereotypes concerning others?

## REFERENCES

1. ADORNO, T. W., FRENKEL-BRUNSWIK, E., LEVINSON, D. S., & SANFORD, R. N. *The Authoritarian Personality*. New York: Harper, 1950.
2. FRENKEL-BRUNSWIK, E. Social research and the problem of values. *J. Abn. & Soc. Psychol.*, 1954, **49**, 466-470.
3. KATZ, D., & BRALY, K. Racial stereotypes of one hundred college students. *J. Abn. & Soc. Psychol.*, 1933, **28**, 280-290.
4. ———. Racial prejudice and racial stereotypes. *J. Abn. & Soc. Psychol.*, 1935, **30**, 175-193.
5. KRECH, D., & CRUTCHFIELD, R. S. *Theory and Problems of Social Psychology*. New York: McGraw-Hill, 1948.
6. MASLING, J. M. How neurotic is the authoritarian? *J. Abn. & Soc. Psychol.*, 1954, **49**, 316-318.
7. SANFORD, F. H. *Authoritarianism and Leadership: A Study of the Follower's Orientation to Authority*. Philadelphia: Institute for Research in Human Relations, 1950.
8. TAYLOR, J. The relationship of anxiety to the conditioned eyelid response. *J. Exp. Psychol.*, 1951, **41**, 81-92.

*VA Hospital*

*Brockton 5, Massachusetts*

## THE EFFECTS OF PRAISE AND REPROOF ON THE GENERALIZATION OF LEARNED CONCEPTS\*

*Department of Psychology, New York University*

---

TATANIA JUZAK<sup>1</sup>

---

### A. THE PROBLEM

The current literature on perception emphasizes selectivity in perceptual organization. It is assumed that this differential selectivity for stimuli is both innate and learned. Studies on human subjects (8, 10, 12) have demonstrated systematic effects of physiological needs, social attitudes, and personal values on perception.

The present study relates to selectivity in generalization of learned concepts. The experiment investigates the effects of praise and reproof during learning of concepts (form and number) upon subsequent preference for these concepts in a situation where generalization of the learned concepts is possible. The literature on perceptual selectivity indicates that motivation states of Ss have been varied by deprivation or threat (1, 8, 11), or reward and punishment for specific responses (13, 14). The independent variables in the present study are types of reward and punishment induced by the use of false norms in the learning situation.

The rationale for the specific hypotheses is derived from principles in learning theory which explain the rôle of reward and punishment: Tolman's principle of emphasis (15), and Hull's principle of reinforcement (7). The specific hypotheses relating to the experimental conditions under which a learned concept will be preferred or non-preferred are:

1. The frequency of preference for a given type of concept learned with praise will be significantly greater than the frequency of preference for the same concept learned with non-praise.

2. The frequency of preference for a given type of concept learned with reproof will be significantly greater than the frequency of preference for the same concept learned with non-reproof.

3. The frequency of preference for a given type of concept learned with

---

\*Received in the Editorial Office on November 29, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This article is an abridgement of a doctoral dissertation submitted to the Graduate School of Arts and Science, New York University. Grateful acknowledgment is due Dr. R. Christie for his guidance.

praise will not differ significantly from the frequency of preference for the same concept learned with reproof.

## B. METHOD

### 1. *Materials*

The stimuli in the eight learning series problems were geometric figures varied in both form and number. The stimuli were drawn in india ink on eight cards, 14" by 14". The sheet was divided into 16 squares: four columns and four rows. In each square were placed figures of different shapes and different number configurations. On four of the problems, a number configuration concept was the correct solution, i.e., similar number configurations were present in each of the four columns. No other similarity of form or number was present in all four columns. On four of the problems, a form concept was the correct solution, i.e., a similar form was present in all four columns.

Four cards, which were similar in size and appearance to the learning series problems, were utilized in the test series. However, each test series

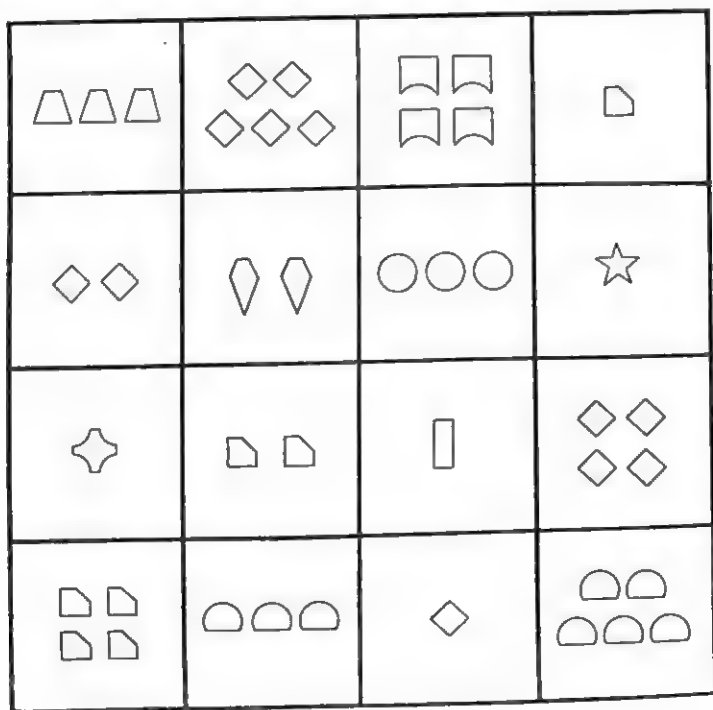


FIGURE 1  
LEARNING SERIES PROBLEM

problem had two correct solutions: one form similarity and one number configuration similarity. None of the forms or number configurations utilized in the test series had been the correct solution during the learning series. Figures 1 and 2 illustrate a learning series problem and a test series problem.

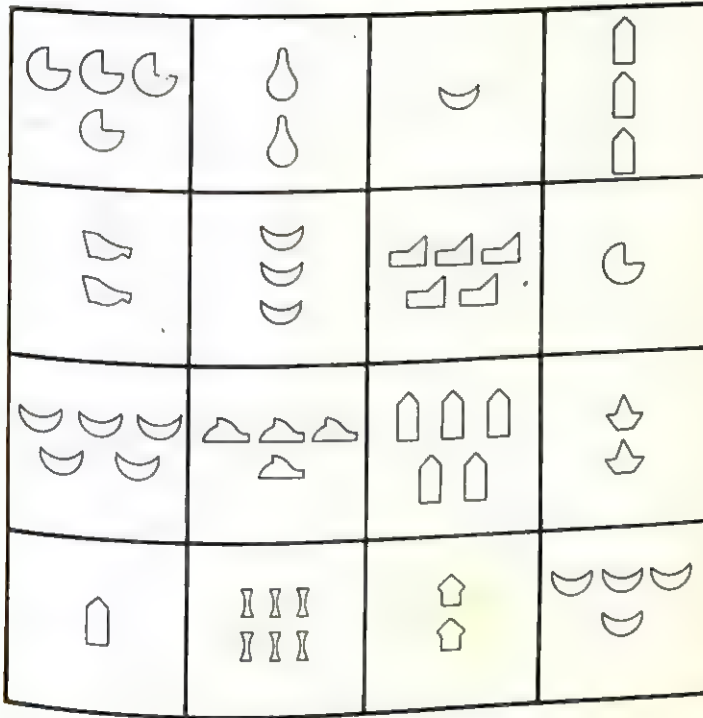


FIGURE 2  
TEST SERIES PROBLEM

The number configurations in both the learning series and the test series problems were constant rather than varied. The results of Heidebreder (4, 5, 6) and Grant (2, 3) demonstrated that the number concept was easier to attain when the number configurations were held constant. Solutions based on constant number configurations and form properties were considered to be comparable in difficulty. This assumption was tested on a sample of 24 Ss. The preference for number or form solutions in the test cards, where two solutions were possible, never exceeded 70 per cent.<sup>2</sup>

<sup>2</sup>In the combined learning and test series problems, 29 different forms were used. However, the available number configurations ranged only from one to seven figures in each unit. Thus, the alternatives of available form and number were not balanced. This problem in method was recognized and resolved on basis of initial experimentation. First, introducing equal availability of form and number alterna-



A final insolvable problem utilized a card similar in size and appearance to the preceding cards.

## 2. Experimental Design

All *Ss* were given the same stimuli cards during the learning and test series. The following differences in experimental treatment were introduced.

*a.* The *P-O* group was praised for the correct solution on one of the two concepts. No praise was given following the correct solution of the second concept.

*b.* The *R-O* group was reproved for the correct solution of one of the two concepts. No reproof was given following the correct solution of the second concept.

*c.* The *P-R* group was praised for the correct solution of one of the two concepts, and reproved for the correct solution of the second concept.

Each of the three groups was composed of two sections in which the praised or reproved concept was varied. Table 1 illustrates the experimental design. Each cell in Table 1 represents a section of a group. In the *NpFo*

TABLE 1  
EXPERIMENTAL DESIGN

Number Form	<i>P-O</i>	Groups <i>R-O</i>	<i>P-R</i>
	<i>NpFo</i> <i>FpNo</i>	<i>NrFo</i> <i>FrNo</i>	<i>NpFr</i> <i>FpNr</i>

section of the *P-O* group, *N* was praised and *F* was non-praised. In the *FpNo* section of the *P-O* group, *F* was praised and *N* was non-praised. The same procedure was followed for the two sections of the *R-O* group except that reproof was substituted for praise. In the *P-R* group both sections received praise and reproof, but the concepts were varied.

## 3. Subjects

The subjects were 96 undergraduate college students: 48 men and 48 women. In each of the six experimental treatments, there were 16 *Ss* equated for sex.

## 4. Experimental Procedure

Each *S* was tested individually. Two sequences of learning series problems were used. The two sequences were: (*a*) *N*, *F*, *N*, *F*, *N*, *F*, *F*, *N*;

tives, i.e., a decrease in form or an increase in number diminished the ambiguous appearance of the problems. Second, the pilot study results indicated that psychological equivalence of form and number concepts was present, despite lack of logical equivalence of alternatives.

(b) F, N, F, N, F, N, N, F. These two sequences were rotated in each group. Prior to testing, the subjects were told that the experiment investigated thinking. The first two problems were practice problems. If the *Ss* failed to attain the solution, it was indicated by *E*. Following the practice problems, the instructions were:

I have norms of reaction times of college students from other universities on this same task. I am going to compare your performance, the length of time it takes you to find the four similar squares, with the performance of these other students. I will tell you in which percentile rank you fall. If the number of the percentile rank is high, such as 86 per cent, it means that you are better than most of these other college students. If the number of the percentile rank is low, such as 13 per cent, it means that you are poorer than most of these other college students. If you get a percentile rank of about 50 per cent, it means that you are average. There will be different norms for each sheet. Remember the directions are the same. Find the four similar squares, point them out and tell the reason why they are similar. Then I shall tell you how your performance for each sheet compares with the performance of other college students.

Subjects were given false norms after solution of each learning problem. The percentile ranks were:

- a. *P-O* group: 75%, 51%, 90%, 47%, 99%, 53%.
- b. *R-O* group: 23%, 51%, 7%, 47%, 2%, 53%.
- c. *P-R* group: 75%, 23%, 90%, 7%, 99%, 2%.

The italicized percentiles in the *P-O* and *R-O* groups represent the praise and reproof percentiles. The percentiles in the *P-O* and *R-O* groups which are not italicized represent the non-praise or non-reproof percentiles. In the *P-R* group, all percentiles represent either praise or reproof percentiles.

*Ss* were given the test series directly following the learning series. The presentation of the four test problems was rotated in four orders of succession in each section of each group. As soon as an *S* gave either of the two solutions to a test problem he was immediately given the next problem. The insolvable problem followed the test problems with no break in procedure. *Ss* were allowed to present solutions to the insolvable problem for five minutes.

The following scores were recorded: (a) Solution time to the nearest second for learning series and test series problems. (b) Preferred solutions on the four test problems. (c) All attempted solutions to the insolvable problem. Following the insolvable problem, each *S* was briefly interviewed.

## C. RESULTS

1. *Test Problems*

No subject demonstrated awareness of the fact that the four test problems had more than one solution. Each subject gave only *one* solution to the four test problems. In the treatment of the results, the percentage frequency of *N* and *F* responses given to all four test problems is used as the basic measure of preference. The percentage frequencies are listed in Table 2.

TABLE 2  
PERCENTAGE FREQUENCIES OF *N* AND *F* RESPONSES  
(*N* is 16 in each section of each group)

	<i>N</i> %	<i>F</i> %	<i>SE</i> %
<i>Total responses</i>			
<i>P-O</i>			
<i>NpFo</i>	26.6	73.4	11.0
<i>FpNo</i>	21.9	78.1	10.3
<i>R-O</i>			
<i>NrFo</i>	50	50	12.5
<i>FrNo</i>	14.1	85.9	8.7
<i>P-R</i>			
<i>NpFr</i>	51.6	48.4	12.4
<i>FpNr</i>	53.1	46.9	12.4
<i>Tests of significance</i>			
	<i>SE<sub>D</sub></i>	<i>D</i>	<i>CR</i>
<i>P-O</i>			
<i>NpFo:FpNo</i>	15.4	4.7	.31
<i>R-O</i>			
<i>NrFo:FrNo</i>	15.2	35.9	2.36*
<i>P-R</i>			
<i>NpFr:FpNr</i>	17.6	1.5	.09

\*Significant at the 5 per cent level of confidence (2.13).

These results which are pertinent to the hypotheses reveal:

a. *P-O* Ss who are *N*-praised give 26.6 per cent *N* responses, and Ss who are non-praised for *N* give 21.9 per cent *N* responses. *F*-praised Ss give 78.1 per cent *F* responses, and Ss non-praised for *F* give 73.4 per cent *F* responses. In the *P-O* group, the frequency of preference for a given type of concept learned with praise is not significantly greater than the frequency of preference for the same concept learned with non-praise. The critical ratio is .31. Hypothesis 1 is not supported.

b. *R-O* Ss who are *N*-reproved give 50 per cent *N* responses, and Ss who are non-reproved for *N* give only 14.1 per cent *N* responses. *F*-reproved Ss

give 86.9 per cent *F* responses and *Ss* non-reproved for *F* give only 50 per cent *F* responses. In the *R-O* group, the frequency of preference for a given type of concept learned with reproof is significantly greater than the frequency of preference for the same concept learned with non-reproof. The critical ratio is 2.36, significant at the 5 per cent level of confidence. Hypothesis 2 is supported.

c. *P-R* *Ss* who are *N*-praised give 51.6 per cent *N* responses, and *Ss* who are *N*-reproved give 53.1 per cent *N* responses. *F*-praised *Ss* give 46.9 per cent *F* responses, and *F*-reproved *Ss* give 48.4 per cent *F* responses. In the *P-R* group, there is no significant difference in frequencies of preference for a given type of concept learned under different conditions of praise and reproof. The critical ratio is .09. Hypothesis 3 is supported.

Further treatment of the data revealed that more *P-R* *Ss* gave both *N* and *F* responses in the four test problems than *Ss* in the *P-O* and *R-O* groups. The data were analyzed for frequency of *N* or *F* "sets" in the three groups. If a subject gave all *N* responses or all *F* responses to the four test problems, he demonstrated a *N* or *F* "set." It was found that *Ss* demonstrated "set" with greater frequency in the *P-O* and *R-O* groups. In the *P-O* group, 62.5 per cent *Ss*, and in the *R-O* group, 71.9 per cent *Ss* gave either four *N* or four *F* responses in the test problems. However, only 40.6 per cent *Ss* in the *P-R* group demonstrated "set." The difference between frequencies of "sets" in the *R-O* and *P-R* groups was statistically significant. The critical ratio was 2.68, significant at the 5 per cent level of confidence.

Additional treatment of this same data revealed an unexpected result: In the *P-O* and *R-O* groups, the frequency of preference for the *F* concept in the four test problems was significantly greater than the frequency of preference for the *N* concept. Following praise for the *F* concept, the frequency of *F* responses (78.1 per cent) was significantly greater than the frequency of *N* responses (26.6 per cent), following praise for *N*. The critical ratio was 3.41, significant at the 1 per cent level of confidence. In the *R-O* group, following reproof for *F*, the frequency of *F* responses (85.9 per cent) was significantly greater than the frequency of *N* responses (50 per cent) following reproof for *N*. The critical ratio was 2.36, significant at the 5 per cent level of confidence.

## 2. Insolvable Problem

The percentage frequencies of *N* and *F* total solutions to the insolvable problem were essentially the same for all three groups. Although the learning situations had been varied in six experimental treatments, chance fre-



quency of *N* and *F* total solutions occurred. For instance, *R-O* *Ss* reprovred for *N* gave 48.3 per cent *N* and 38.5 per cent *F* total solutions.<sup>3</sup> *R-O* *Ss* reprovred for *F* responded with 49.1 per cent *N* and 38.5 per cent *F* total solutions. These percentage frequencies of the *R-O* group were similar to the percentage frequencies obtained for the two sections of the *P-O* and *P-R* groups.

The percentage frequencies of *N* and *F* initial solutions to the insolvable problem are listed in Table 3. There are no significant differences in per-

TABLE 3  
PERCENTAGE FREQUENCIES OF *N* AND *F* SOLUTIONS TO INSOLVABLE PROBLEM  
(*N* is 16 in each section of each group)

	<i>N</i> solutions %	<i>F</i> solutions %	<i>SE</i> %
<i>Initial Solutions</i>			
<i>P-O</i>			
<i>NpFo</i>	37.5	62.5	12.1
<i>FpNo</i>	31.2	68.8	11.5
<i>R-O</i>			
<i>NrFo</i>	56.2	43.8	12.4
<i>FrNo</i>	12.5	87.5	8.2
<i>P-R</i>			
<i>NpFr</i>	37.5	62.5	12.1
<i>FpNr</i>	37.5	62.5	12.1
<i>Tests of Significance</i>			
	<i>SE<sub>D</sub></i>	<i>D</i>	<i>CR</i>
<i>P-O</i>			
<i>NpFo:FpNo</i>	16.7	6.2	.37
<i>R-O</i>			
<i>NrFo:FrNo</i>	14.9	43.7	2.93**
<i>P-R</i>			
<i>NpFr:FpNr</i>	17.1	0.0	0.0

\*\*Significant at the 1 per cent level of confidence (2.92).

centage frequencies of *N* or *F* initial solutions in the two sections of the *P-O* and *P-R* groups. However, there is a significant difference in *N* or *F* initial solution frequencies in the *R-O* group. *Ss* reprovred for a concept during learning apply that concept with significantly greater frequency in the initial solution to the insolvable problem than *Ss* non-reprovred for the same concept.

<sup>3</sup>Another category of solutions not listed was 0 solutions (solutions other than *N* or *F*).

## D. DISCUSSION

The results indicated that reproof, or praise combined with reproof, during learning of concepts influenced preference for concepts in a subsequent generalization situation. With the condition of reproof, the reproofed concept was preferred. With the condition of combined praise and reproof, the praised and reproofed concepts were relatively equally preferred. However, praise alone during learning did not influence preference for the praised concept. The ineffectiveness of praise alone during learning is explained on the basis of absence of drive conditions aroused by the use of false norms. This drive condition is described by McClelland (9) as a need for achievement. McClelland defines "need achievement" as the learned drive which is aroused by failure and satisfied by success.

In both the *R-O* and *P-R* groups, where either low scores, or high scores combined with low scores were presented, subjects were challenged to overcome the failure represented by the low scores. The need to achieve was aroused. *R-O* Ss acquired a readiness for the reproofed concept because the reproofed concept represented failure which must be overcome. On this basis, the reproofed concept became more salient than the same concept non-reproofed. *P-R* Ss acquired readiness for both praised and reproofed concepts, because the reproofed concept represented failure which was challenging and the praise concept represented the satisfaction of success. However, in the *P-O* group, praise alone during learning presented no failure situation. The need to achieve was not aroused because *P-O* Ss were given only high and neutral scores. Thus, *P-O* Ss were not motivated by praise. No specific effects determined by praise were demonstrated in subsequent preference for concepts.

The results indicated an unexpected dominant preference for the *F* concept in the *P-O* and *R-O* groups. Preference for the *N* concept never exceeded preference for the *F* concept even when praise or reproof were presented to accentuate the *N* concept over the *F* concept. It was found that *N* and *F* preferences were similar following learning conditions in which praise and reproof were absent. However, a dominant form preference occurred following learning with praise or reproof. Such results suggested that a bias relating to differences in eliciting *N* and *F* responses was present when praise and reproof were introduced, although this bias was not present following neutral learning. This bias is explained on the basis of differences in accessibility of the *N* and *F* concepts which occurred when praise

and reproof were introduced and which was absent following neutral learning.

The studies of Heidbreder (4, 5, 6) and Grant (2, 3) offer evidence that accessibility of *N* and *F* concepts may differ depending upon differences in stimulus presentations. Heidbreder has indicated that the form concept is more perceptually accessible than the number concept because the properties which identify the form concept are perceptually apparent, and the properties which identify the number concept are abstract and not perceptually apparent. Grant has indicated that two responses can be associated with number: a perceptual response when the stimulus configuration of the number of figures is kept constant, and an analytic response when the stimulus configuration is varied. On the basis of Heidbreder's and Grant's findings the test material was arranged so that both number and form properties were perceptually apparent. Yet, in the *P-O* and *R-O* groups, despite this equivalence in situational support, the readiness to perceive form properties was much stronger than the readiness to perceive number properties. The explanation which is offered to account for the differences in accessibility of *N* and *F* concepts is that, despite equivalence in situational support, there occurred differences in the types of responses by which *N* and *F* properties were discriminated.

The *F* concept was more perceptually accessible than the *N* concept because only one response, a perceptual response, could be made to form properties. The *N* concept was not always as perceptually accessible as the *F* concept because two responses, either a perceptual or an analytic response could be made to *N* properties. In the *N* practice problem, *Ss* did not offer a readily available perceptual response, since the *Ss* did not perceive number similarity. When the four numerically similar squares were indicated to them, a frequent spontaneous comment to this *N* practice problem was "I never thought of counting." The initial association to *N* was a counting or analytic response. There is evidence which indicates that, during the insolvable problem, the frequencies of *N* and *F* total solutions were relatively equal in all three groups, thereby suggesting that *N* and *F* concepts were equally accessible. However, during the test series, *N* and *F* frequencies differed with differential motivation. It is suggested that the accessibility of the *N* concept was weakened because there was a possibility for two types of responses to *N* properties.

In the *R-O* group, the accessibility of the *N* concept varied with reproof. If *N* was reproofed, both *N* and *F* were equally accessible in the test problems. If *N* was non-reproofed, the accessibility of *N* was significantly dimin-

ished, and *Ss* showed an overwhelming preference for *F*. In the *P-R* group, *N* and *F* were equally accessible, whether *N* was praised or reproved. In the *P-O* group, praise for *N* was ineffective in increasing the accessibility of *N*. *P-O Ss* showed significantly greater preference for *F*, independent of which concept was praised.

Previous studies have demonstrated selectivity in perception. The present study presents data demonstrating selectivity in generalization of learned concepts.

### E. SUMMARY

This study investigated the effects of praise and reproof during learning of two concepts (form and number) upon subsequent preference for these concepts in a situation where generalization of the learned concepts was possible. Praise and reproof were presented by use of false norms. There were three experimental groups: a group where concepts were praised and non-praised during learning; a group where concepts were reproved and non-reproved during learning; a group where concepts were both praised and reproved during learning. The results indicated:

1. The frequency of preference for a concept learned with praise was not significantly greater than the frequency of preference for the same concept learned with non-praise.

2. The frequency of preference for a concept learned with reproof was significantly greater than the frequency of preference for the same concept learned with non-reproof.

3. The frequency of preference for a concept learned with praise did not differ significantly from the frequency of preference for the same concept learned with reproof.

4. In the two groups, where concepts were praised and non-praised, or reproved and non-reproved, the form concept was preferred with significantly greater frequency than the number concept.

The drive conditions induced by praise and reproof were described as a need for achievement. It was suggested that praise alone during learning did not arouse the need to achieve. The dominant form preference in two groups was explained on the basis of differences in accessibility of the *N* and *F* concepts.

### REFERENCES

1. BRUNER, J. S., & POSTMAN, L. Tension and tension release as organizing factors in perception. *J. Personal.*, 1947, **15**, 300-308.
2. GRANT, D. A. Perceptual versus analytical responses to the number concept of a Weigl-type card sorting test. *J. Exp. Psychol.*, 1951, **41**, 23-29.



3. GRANT, D. A., JONES, O. R., & TALLANTIS, B. The relative difficulty of the number, form, and color concepts of a Weigl-type problem. *J. Exp. Psychol.*, 1949, **39**, 552-557.
4. HEIDBREDER, E. The attainment of concepts: I. Terminology and methodology. *J. Gen. Psychol.*, 1946, **35**, 173-189.
5. ———. The attainment of concepts: II. The problem. *J. Gen. Psychol.*, 1946, **35**, 191-223.
6. ———. The attainment of concepts: III. The process. *J. of Psychol.*, 1947, **24**, 93-138.
7. HULL, C. L. Principles of Behavior. New York: D. Appleton Century, 1943.
8. MCCLELLAND, D. C., & ATKINSON, J. W. The projective expression of needs: I. The effect of different intensities of the hunger drive on perception. *J. of Psychol.*, 1948, **25**, 205-222.
9. MCCLELLAND, D. C., CLARK, R. A., ROBY, T. B., & ATKINSON, J. W. The projective expression of needs: IV. The effect of the need for achievement and thematic apperception. *J. Exp. Psychol.*, 1949, **39**, 242-255.
10. MCCLELLAND, D. C., & LIBERMAN, A. M. The effect of need for achievement on recognition of need-related words. *J. Personal.*, 1949, **18**, 236-251.
11. POSTMAN, L., & BRUNER, S. Perception under stress. *Psychol. Rev.*, 1948, **55**, 314-323.
12. POSTMAN, L., BRUNER, J. S., & MCGINNIES, E. Personal values as selective factors in perception. *J. Abn. & Soc. Psychol.*, 1948, **43**, 142-154.
13. PROSHANSKY, N., & MURPHY, G. The effects of reward and punishment on perception. *J. of Psychol.*, 1942, **13**, 295-305.
14. SHAFER, R., & MURPHY, G. The rôle of autism in a visual figure-ground relationship. *J. Exp. Psychol.*, 1943, **32**, 335-343.
15. TOLMAN, E. C., HALL, C. S., & BRETNALL, E. P. A disproof of the law of effect and a substitution of the laws of emphasis, motivation, and disruption. *J. Exp. Psychol.*, 1932, **15**, 601-614.

*Bellevue Psychiatric Hospital*  
*104 West 90th Street*  
*New York City 24*

## STIMULUS-DISCRIMINABILITY AND CONCEPT-ATTAINMENT: A QUESTION ARISING FROM BAUM'S EXPERIMENT\*

*Department of Psychology, Wellesley College*

---

EDNA HEIDBREDER

---

### A. INTRODUCTION

Baum has reported an experiment (1) in which she repeated, with specified modifications in technique, one of my experiments on the attainment of concepts, Experiment *B* (9). Her experiment was performed to discover whether the results obtained in *B* might not be interpreted in terms of Gibson's hypothesis (5) on the rôle of stimulus-discriminability<sup>1</sup> in verbal learning;—more specifically, whether the attainment of concepts in *B* might not be interpreted as a special case of paired-associate learning, and the order of their attainment as a function of the discriminability of the presented stimulus items. According to my interpretation, the concepts in *B* were attained more or less rapidly as the critical features<sup>2</sup> of their respective instances—indicated in *B* by drawings—were more or less thing-like. This interpretation was made in terms of a hypothesis which has been presented elsewhere (8). That part of it which is most relevant to the present discussion proposes an order of dominance among human cognitive processes such that the perception of concrete objects is predominant among them, and such that other cognitive processes are decreasingly dominant as they involve increasingly great variations from or modifications of the kind of process involved in perceiving concrete objects.

The results of Baum's experiment, which will be called *B'* in this paper, conformed closely to those obtained in *B*. In both experiments, concepts of objects, forms, and numbers were attained in that order and in *B'*, as in *B*, the order was statistically well established. This confirmation is itself worthy of notice. However, on the basis of data concerning first-list learning, Baum interpreted the results as supporting Gibson's discriminability hypothesis.

---

\*Received in the Editorial Office on December 4, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>The terms 'stimulus-discriminability' and 'generalization,' and also such terms as 'stimulus,' 'stimulus item,' and 'list' will be used in this paper as they are used by Gibson (5, 6, 7) and Baum (1, 2).

<sup>2</sup>The critical feature of an instance is that feature which is present in all instances of a concept and not present in non-instances of that concept.

In  $B'$  the conditions of  $B$  were reproduced with great fidelity. In general, the changes in technique were such as to secure a stricter control than in  $B$  of some of the experimental conditions, especially those relevant to the learning that occurs during such experiments. An especially important improvement introduced into  $B'$  was that of recording the  $S$ s' responses not only, as in  $B$ , on a protocol sheet, but also on a magnetic tape. In  $B$  the record indicated only whether a response was correct or incorrect, but in  $B'$  the tape recordings made it possible to supplement the record of correct and incorrect responses by a detailed analysis of the *kinds* of errors that occurred in  $B'$ . Among these was a class of errors called *intrusions*—a kind of error in which  $S$  responds to a presented drawing with a syllable learned as the name for another drawing in the same experiment.<sup>3</sup> Such errors were assumed to indicate that the two drawings had not been adequately discriminated, and mean intrusions per drawing—called intrusion-scores in the present paper—were used as measures of stimulus-discriminability. So interpreted, intrusion-scores were positively correlated with the order in which correct responses were acquired during first-list learning, and also with the order in which the concepts were eventually attained. Largely on the basis of this evidence, the results of  $B'$  were presented as supporting the discriminability hypothesis.

The data from  $B'$  were analyzed and interpreted on the assumption that the two hypotheses under consideration are alternative in the sense that the acceptance of one makes the other unnecessary. This assumption, though not explicitly stated, is clearly implied, especially in the statement that a choice between the two is in order (1).<sup>4</sup> Several points in the interpretation of  $B'$  are open to question (among them the assumption that the two hypotheses are alternative) but the present paper will be concerned with a single question: whether intrusion-scores, which supply the measure basic to Baum's interpretation, are acceptable, as used in  $B'$ , as valid measures of stimulus-discriminability.

#### B. THE DATA AND BAUM'S INTERPRETATION

The new factual evidence provided by  $B'$  was based on a detailed analysis of first-list responses. These consisted of correct responses and three classes of errors: omissions, invented syllables, and intrusions.

<sup>3</sup>This definition was based on Gibson's definition of 'generalization' as "the tendency of a response  $R_a$  learned to  $S_a$  to occur when  $S_b$  (with which it has not been previously associated) is presented" (5, p. 204).

<sup>4</sup>Baum cites experiments by Dattman and Israel (3) as weakening the case for my hypothesis. My interpretation of their experiments has been presented elsewhere (10).

The most numerous of the errors were omissions, i.e., failures to name the presented drawing at all. In the paper under discussion, no data were presented concerning omissions, perhaps because the conditions did not provide an adequate basis for sorting out, among the many possibilities, the probable determinants of such failures to respond. From a fuller report of *B'* (2) however, it is possible to calculate that well over half of the first-list errors—about 63 per cent—were omissions. Omissions were increasingly frequent for object-drawings, form-drawings, and number-drawings,<sup>5</sup> thus showing a definite and consistent pattern of relationship both with the order in which correct first-list responses were learned, and with the order in which the corresponding concepts were attained.

Baum's analysis, however, was concerned chiefly with *overt* errors. Of these, the more numerous were those called invented syllables, i.e., responses in which *Ss* produced nonsense syllables not among those assigned as names of drawings in this experiment. They made up about 30 per cent of the first-list errors. As compared with other measured behavior in *B'* they formed a class apart. They were significantly correlated neither with the order of concept-attainment ( $r = -.12$ ) nor with number of correct first-list responses ( $r = .10$ ) nor with number of intrusions ( $r = -.23$ ). In frequency they did not vary significantly from drawing to drawing, and the slight variations that did appear followed no consistent direction or pattern.

As compared with invented syllables, the other overt errors, intrusions (defined above) present a very different picture. Making up the remaining 7 per cent of first-list errors, they were increasingly frequent for object-drawings, form-drawings, and number-drawings in that order, and they were significantly correlated, as the other overt errors were not, both with the order of concept-attainment ( $r = .69$ ;  $p = < .05$ ); and with the frequency of correct responses per drawing during first-list learning ( $r = -.86$ ;  $p = < .01$ ). On the assumption that first-list intrusion-scores measure stimulus-discriminability, and because intrusions were the only overt errors significantly correlated with first-list learning, Baum interpreted these data as indicating that the concepts in *B'* were attained more or less rapidly as first-list learning occurred more or less rapidly for the drawings indicating their instances; and that such learning occurred more or less rapidly as the drawings were more or less discriminable.

<sup>5</sup>The terms 'object-drawings,' 'form-drawings,' and 'number-drawings' are abbreviations, used to refer to drawings which indicate instances, the critical features of which are respectively concrete objects, spatial forms, and numerical quantities. It is to be understood that all the indicated instances included non-critical as well as critical features.



In sum, the conclusion was that intrusion scores provide the clue to the rates of first-list learning per drawing in  $B'$ , and that the data on first-list learning support the discriminability hypothesis as an interpretation of the order of concept-attainment in that experiment.

### C. ANOTHER VIEW OF THE DATA

But when *all* measures of first-list learning are viewed together and as a whole, when no one of them is selected as a special center of interest, it becomes apparent that, with the single exception of invented syllables, they all fall consistently into a single, inclusive pattern. All of them—correct responses, intrusions, and omissions—indicate that the  $S$ s learned to respond correctly to object-drawings, form-drawings, and number-drawings in approximately that order, and thus in approximately the order in which the concepts were eventually attained. It is especially noteworthy that intrusions and omissions follow the same pattern in their relationships both to other measures of behavior and to the stimulus-drawings. The measure that stands apart is not that of intrusions, but that of invented syllables, and this measure is clearly related to the acquisition of the correct verbal responses rather than to the identification of the drawings. The other measures, all consistently correlated with each other, with the order of concept-attainment, and with the kinds of drawings responded to, form a single, total pattern in which the component measures together suggest that the factors that determined the eventual order of concept-attainment were already operative during first-list learning, and that those factors were related to the identification of the drawings rather than to the acquisition of the verbal responses.

Seen in this context, intrusions appear as a part—and numerically a very slight part—of a large, inclusive pattern. Since they are not unique<sup>6</sup> in their correlation with the order of first-list learning and of concept-attainment, it cannot be taken for granted that they supply the sole, or the chief, or even an especially important clue to the order of either of those achievements. They may do so, but the data are also consistent with other interpretations—among them that correct first-list responses were *learned* more or less rapidly as the critical features of the indicated instances were more or less thing-like. This relationship was clearly indicated in  $B'$ , as it was not in  $B$ , where it may have been obscured by a less adequate control of the con-

<sup>6</sup>Baum reports that "only the intrusion errors are significantly correlated with correct responses [during first-list learning]" (1, p. 92). Here the reference is evidently to *overt* errors, which include invented syllables and intrusions, but not omissions.

ditions of first-list learning than that secured in  $B'$ .<sup>7</sup> In other words, the data of  $B'$  may indicate that the factors said to be determinants of the order of concept-attainment in  $B$ , were actually at work earlier than was apparent from the data of  $B$ .

Furthermore, intrusions themselves may be interpreted in a manner not considered in the report of  $B'$ . There the high negative correlation between first-list intrusions and first-list correct responses per drawing was presented as evidence that first-list learning was determined chiefly by the discriminability of the drawings. But this correlation may indicate, instead, that as  $S$  learns to name more and more drawings correctly, the chances for intrusions become increasingly great for the remaining drawings and names.

During first-list learning, an important part of  $S$ 's task consists in memorizing nonsense-syllables, i.e., in acquiring the small stock of verbal responses which includes *all* the responses and *only* the responses accepted as correct. By definition, intrusions are members of this small, closed class of acceptable responses. Consequently, with every drawing  $S$  learns to name correctly, there is a corresponding decrease in the number both of the drawings to be named and of the acceptable responses for naming them, and therefore a corresponding increase in the chances that just those remaining responses will be misapplied to just those remaining drawings. Thus *whatever* the determinants of the order in which correct responses are learned, chance would favor a distribution of intrusions in which their frequency per drawing varied inversely with the frequency of correct responses per drawing, and thus directly with the order in which correct responses were learned: in other words, the distribution obtained in  $B'$ . The distribution of intrusions, therefore, might be a consequence of the order in which correct responses were learned, without itself giving any direct information about the determinants of that order.

Other factors, operating singly or together, might also occasion intrusions in  $B'$ —among them guessing, momentary fluctuations in the availability of acceptable verbal responses, and interferences among such responses, few in number, all kept in readiness by practice and by approximately equal amounts of practice.

Because intrusions may be occasioned in many ways, it becomes necessary to ask on what grounds intrusion scores were accepted in  $B'$  as measures of stimulus-discriminability.

<sup>7</sup>A more adequate control was secured in  $B'$  by appropriate variations (a) in pairings of drawings and names, and (b) in the order of presentation of first-list items.

## D. INTRUSION SCORES AND STIMULUS-DISCRIMINABILITY

This measure, as noted above, was suggested by Gibson's hypothesis, and Baum cites evidence from several experiments as justifying its use in actual research (4, 6, 7, 11, 12).

But intrusions are *behavior*; measures based on intrusions are primarily *measures of behavior*; and the only direct information such measures give about *stimuli* is that in their presence, and in the given conditions, such behavior occurred to the measured degree. It is *stimuli*, however, which are said to be more or less discriminable. How is the transition made from a measure of behavior to a description of stimuli? With what characteristic of stimuli, as distinguished from responses, is the measured behavior said to be correlated?

According to Gibson's hypothesis and experimental practice, with their *similarity* to each other. In Gibson's words, "Similarity will be considered . . . as *that relationship among stimulus items* (ital. mine) which can be indicated and measured in terms of their *tendency to generalize* (ital. Gibson's)" (5, p. 208). Here a distinction seems to be made—though this is by no means certain—between *indiscriminability* as a characteristic of stimuli which is inferrable from a specified kind of behavior, and *similarity* as a characteristic of stimuli which can be referred to independently of that behavior, and which is presumably not wholly reducible to it. At least the similarity among stimuli is not treated as describable solely in terms of their indiscriminability; illustrations are added to indicate in what respects, along what dimensions or continua, stimuli may vary in similarity—"e.g., a series of tones, or of spots along the foreleg" (5, p. 208). In the initial statement of the hypothesis, this problem receives little attention, but the suggested distinction becomes evident and important when experiments are performed to test the hypothesis.

In all the experiments cited by Baum as relevant to this point, a distinction was observed in practice, if not in theory, between the *similarity* of the stimulus materials as introduced into the experiment, and their *indiscriminability* as measured by behavior occurring during the experiment. In all of them, the stimulus-materials were pre-experimentally prepared as similar, and in all of them the criteria of similarity, which were applied pre-experimentally, were both *independent of* and *different from* the indicators of indiscriminability (intrusions) used in the experiment proper. Among the criteria of similarity were physical measurements, judges' ratings, and the construction of drawings as 'variations' from a standard drawing—all involving forms

of behavior clearly different from that involved in intrusions.<sup>8</sup> In other words, 'similarity' and 'indiscriminability' as applied to the stimulus-materials, are not synonymous terms, as operationally defined. If not in the stated hypothesis, at least in the experimental practice associated with it, the use of intrusions as indicators of indiscriminability presupposes that those materials are similar in some respect which is specifiable in terms independent of those derived from the experimental data themselves.

In *B'* however, no dimension or continuum was specified along which the stimulus-items vary in similarity. On the contrary, intrusions were used as a means of securing information about the stimulus-drawings themselves. These were said to be more or less discriminable as intrusion-scores were low or high. Yet in none of the experiments cited is there justification for the use of intrusions in this way. All of those experiments began with stimulus-materials *introduced* into the experiment as similar, and all of them proceeded to investigate various aspects of learning, for the most part certain transfer effects, some of which were indicated and measured by intrusions errors. What those experiments show is that *when* stimulus materials are similar, and *when* conditions are such as to make 'correct' responses dependent on discriminations among them, certain predicted transfer effects, including intrusions, occur. From this, of course, it does not follow that intrusions, by their mere occurrence, give evidence of stimulus-indiscriminability, whatever the conditions in which they occur.

Without *some* specifiable similarity among the stimulus-items, intrusion-scores, which supply the data basic to the interpretation of *B'*, cannot be accepted as measures of stimulus-discriminability. Indeed, without *some* description of the stimulus-drawings in terms not derived from the experimental data, no relationship can be specified between the drawings and the data; i.e., no relationship can be specified between intrusion-scores and dependent variables, and the reported correlations reported in *B'*—are simply other measures—and in fact all the correlations reported in *B'*—are simply correlations among dependent variables. It is significant that, within the framework of *B'*, where it is essential that the drawings be characterized

<sup>8</sup>In Gibson's own experiments, the pre-experimental preparation of the stimulus-materials was especially thorough. It included a step which so closely resembles intrusions that it is perhaps not an exaggeration to say that the materials were pre-experimentally tested for indiscriminability, not merely for similarity. But this step was only a single, final step in a process of preparation which included the construction of 'variations' from standard drawings and judges' ratings of these, and thus the use of criteria of similarity which differed from the measure of indiscriminability used in the experiments themselves (6, 7).



as more or less discriminable, it is not possible to answer the question: discriminable in what respect?

One way of meeting this difficulty might be to accept degrees of thing-character as constituting a dimension or continuum along which the stimulus-items vary in similarity. To do so would have the additional advantage of taking into account a clearly indicated relationship not included in Baum's interpretation: the consistent pattern of correlations between the various measures of first-list learning (with the exception noted above) and the order: object-drawings, form-drawings, number-drawings. But to adopt this expedient, whether implicitly or explicitly, would be to make degrees of thing-character a condition of the degrees of discriminability said to be indicated in *B'*, and thus to make an interpretation in terms of stimulus-discriminability supplementary to—not alternative to and substitutable for—an interpretation which includes degrees of thing-character among the experimental conditions and among the situational determinants of the *Ss'* behavior. It is possible, of course, that some dimension other than thing-character may be found, along which the stimulus-drawings vary in similarity as distinguished from indiscriminability. But no such dimension has been specified for the drawings used in *B'*. The experiment, therefore, fails to meet the conditions required by the discriminability hypothesis, at least as it is employed in actual experimental research, for the use of intrusions as indicators of stimulus-indiscriminability.

### E. TWO COMMENTS

It may be well, to prevent misinterpretations, to add the following comments.

The first concerns Baum's assumption that the two hypotheses in question are alternative in the sense previously noted. The position taken in the present paper is that they are not alternative. It is assumed in my hypothesis that learning is "importantly involved" (8, p. 2) both in the perception of objects and in the attainment of concepts. But no assumption is made concerning the specific nature of learning; on the contrary the hypothesis is concerned with a pattern of molar relationships and is independent of any particular theory of learning. It is unnecessary in this hypothesis either to accept or to reject any particular theory of how learning occurs. Any evidence therefore that a specified kind of learning is involved in concept-attainment may supplement the hypothesis without constituting an alternative interpretation. Perhaps it should also be noted that even if it had been established in *B'* that object-drawings, form-drawings, and number-drawings

were decreasingly discriminable in that order, a question would still arise as to whether the indiscriminability hypothesis is sufficient in and by itself to account for that correlation, which is basic to the interpretation of the order of concept-attainment in *B* and *B'* in terms of the discriminability hypothesis.

The second comment is that no denial is made in this paper that intrusions may, and sometimes do, indicate stimulus-indiscriminability in experiments like *B'*. Indeed, responses of the sort called intrusions in *B'* have not only been reported in my papers, but have been interpreted, when supplemented by other evidence, as indicating that the *Ss* had confused some of the drawings, i.e., in the terms of the discriminability hypothesis, had not adequately discriminated them.<sup>9</sup> As a matter of fact, the indicated confusions—since they occurred *before* the critical features had been identified, and since they involved non-critical features which were *more* thing-like than the critical features of the drawings in question—were presented as part of the evidence that the more thing-like features of a drawing had a consistent advantage over the less thing-like in becoming the dominant situational determinants of overt naming responses. The point made in this paper is that intrusions, though they may indicate that stimuli are indiscriminable do not always do so; and that in *B'* the conditions have not been met which make intrusion-scores acceptable as valid measures of stimulus-discriminability.

#### F. SUMMARY

Intrusion scores, as used in *B'*, are not acceptable as valid measures of stimulus-discriminability. The data of *B'* therefore give no support to the indiscriminability hypothesis as an interpretation of the order of concept-attainment in that experiment.

Intrusion scores, as used in *B'*, fail to qualify as valid measures of stimulus-discriminability, partly because factors other than stimulus-indiscriminability may have occasioned intrusions in *B'*, chiefly because no dimension has been specified along which the stimulus-items can be ordered with reference to similarity, as distinguished from indiscriminability. If degrees of thing-character are accepted implicitly or explicitly as constituting such a dimension, the discriminability hypothesis becomes supplementary to—not alternative to and substitutable for—an interpretation in which degrees of thing-character are recognized among the experimental conditions and among the determinants of the *Ss'* behavior.

<sup>9</sup>For example, in Experiments *A*, *B*, and *C*, which most closely resemble *B'*, such responses have been noted (9, pp. 113-14, 118, 128).

## REFERENCES

1. BAUM, M. H. Simple concept learning as a function of intralist generalization. *J. Exp. Psychol.*, 1954, **47**, 89-94.
2. ———. A study in concept attainment and verbal learning. Unpublished doctoral dissertation, Yale University, 1951.
3. DATTMAN, P., & ISRAEL, H. The order of dominance among conceptual capacities: An experimental test of Heidebreder's hypothesis. *J. of Psychol.*, 1951, **31**, 147-160.
4. GAGNÉ, R. M. The effect of sequence of presentation of similar items on the learning of paired associates. *J. Exp. Psychol.*, 1950, **40**, 61-73.
5. GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196-229.
6. ———. Retroactive inhibition as a function of degree of generalization between tasks. *J. Exp. Psychol.*, 1941, **28**, 93-115.
7. ———. Intra-list generalization as a factor in verbal learning. *J. Exp. Psychol.*, 1942, **30**, 185-200.
8. HEIDBREDER, E. Toward a dynamic psychology of cognition. *Psychol. Rev.*, 1945, **52**, 1-22.
9. ———. The attainment of concepts: III. The process. *J. of Psychol.*, 1947, **24**, 93-138.
10. ———. Experiments by Dattman and Israel on the attainment of concepts. *J. of Psychol.*, 1952, **34**, 115-136.
11. NEWMAN, S. E. The effects of similarity and contiguity on the formation of concepts. Unpublished doctoral dissertation, Northwestern Univ., 1951.
12. OSEAS, L., & UNDERWOOD, B. J. Studies of distributed practice: V. Learning and retention of concepts. *J. Exp. Psychol.*, 1952, **43**, 143-148.

*Department of Psychology*  
*Wellesley College*  
*Wellesley, Massachusetts*

## DRAWINGS OF A THREE-DIMENSIONAL OBJECT BY MENTAL PATIENTS: A PRELIMINARY REPORT\*

*Department of Psychology, Miami University*

C. W. CRANNELL AND ERIKA PLAUT<sup>1</sup>

The purpose of the writers is to report upon a phase of research in progress on the evaluation of drawings made by schizophrenics. The stimulus for their research was provided by two studies: Vernier (3) and Whitmyre (4). Vernier presents a wide variety of drawings of the human form, male and female, by psychotics, psychoneurotics, and normals. Her material is purposely of a selected type, and is intended as a kind of handbook of projective drawings. Each drawing is accompanied by an adequate case history and an analysis of the drawings. Whitmyre's study offers a critical challenge to the use of such materials. He found that drawings of the human form by abnormals were judged significantly only in terms of their artistic excellence, and that persons well trained in this projective technique were not only unable to discriminate between normal and abnormal human form drawings, but furthermore, when these experts rated drawings on artistic excellence, they produced evaluations which showed substantial correlation with their judgments of abnormality.

Whitmyre's study once more arouses the suspicion of "contamination" in the utilization of this projective technique. If the technician knows the psychiatric analysis, is it perhaps that he "reads into" the drawing certain peculiarities which he feels are representative of the particular syndrome? Are these signs "seen" in the drawing no different from the same aspects of drawings by poorly skilled normals?

The present writers do not believe that these issues are likely to be solved as long as research on projective drawings is restricted, as seems to be the case in recent years, to drawings of the human form. Kendig points out, in her introduction to Vernier's book, that the subject's identification with the figure drawn is often frankly acknowledged, and that "there seems to be considerable safety in inferring at least unconscious identification in the majority of other cases" (3, p. vii). It seems possible that asking the patient

\*Received in the Editorial Office on December 5, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.  
<sup>1</sup>The writers wish to thank Dr. J. A. Mendelson, Superintendent, and Mr. Scott Bowers, Chief Psychologist, of the Dayton State Hospital for all the kind assistance which has made this research possible.



to make a drawing of some object with which identification need not be so direct as that with the human form may well produce results less inhibited than those of the human form so often appear to be.

The object used by the writers is a large silver elephant, illustrated in an earlier article by the senior writer (1). The choice of this object was based upon the fact that the writers have available over 350 drawings by normals (college students), both with and without art training. This collection of drawings has been subjected to many evaluations, and it appeared reasonable to consider this material as a background against which drawings by groups of abnormals could be contrasted.

The drawings by patients were made in the hospital ward. Directly after the drawing was completed, the Shipley-Institute of Living (Hartford Retreat) Scale for Measuring Mental Impairment (2) was administered. Fifty patients, all but three diagnosed as schizophrenic, were approached with these two tasks; all but seven of them made a definite effort to respond, and the majority of them appeared to be trying quite hard to please the examiner.

The drawings chosen for presentation in this preliminary report are highly selected. Five of them were drawn by hospitalized patients and one by a normal college student. None is necessarily typical of the group from which it was drawn. In Figure 1 the six drawings are lettered *A* to *F*. The case histories now to be presented are numbered from 1 to 6, but not in the same order as the letters. The reader is invited to attempt matching the drawings with the case histories before checking the key.

*Case 1.* Female, age 45, 8th grade, Shipley Scale Vocabulary Age (*SVA*) 10.3, Shipley Conceptual Quotient (*CQ*) 100, Wechsler Full Scale 91 (61 mo. before drawing), hospitalization 62 mo., paranoid schizophrenia. Good childhood adjustment. Married at age 16 and divorced after one year; remarried at 23. Extramarital love affair for several years followed by divorce; abandoned by extramarital lover. Became unable to hold job for more than few days; ran up excessive bills; developed delusions parents were trying to kill her.

*Case 2.* Female, age 38, 7th grade, *SVA* 10.3, *CQ* 86, no Wechsler, hosp. 52 mo., schizophrenia with hebephrenic, catatonic and paranoid symptoms. Moody as child. Unhappy because parents broke up her engagement; jealous of married sisters. Became irresponsible at age 34; taken to hospital after auto wreck. Has temper tantrums and destructive periods amplified at menstruation.

*Case 3.* Female, age 28, 8th grade, *SVA* 11.9 (but language handicap), *CQ* 74, Wechsler 80 (4 mo. before drawing), hosp. 18 mo., paranoid schizophrenia. Latin American. Mother died when patient was 5, reared by relatives. Marriage at 20 opposed by both families be-

cause of social difference. After second child refused intercourse with husband for fear of pregnancy. Delusions of grandeur; delusions of poisoning. Attacked husband; had to be restrained from running nude in public. Before illness was timid, affectionate, insecure.

*Case 4.* Female, age 34, 9th grade, *SFA* 13.5, *CQ* 73, no Wechsler, hosp. 116 mo.; diagnosis on admission, psychopathic personality without psychosis, 5 mo. later, paranoid schizophrenia. Problem child

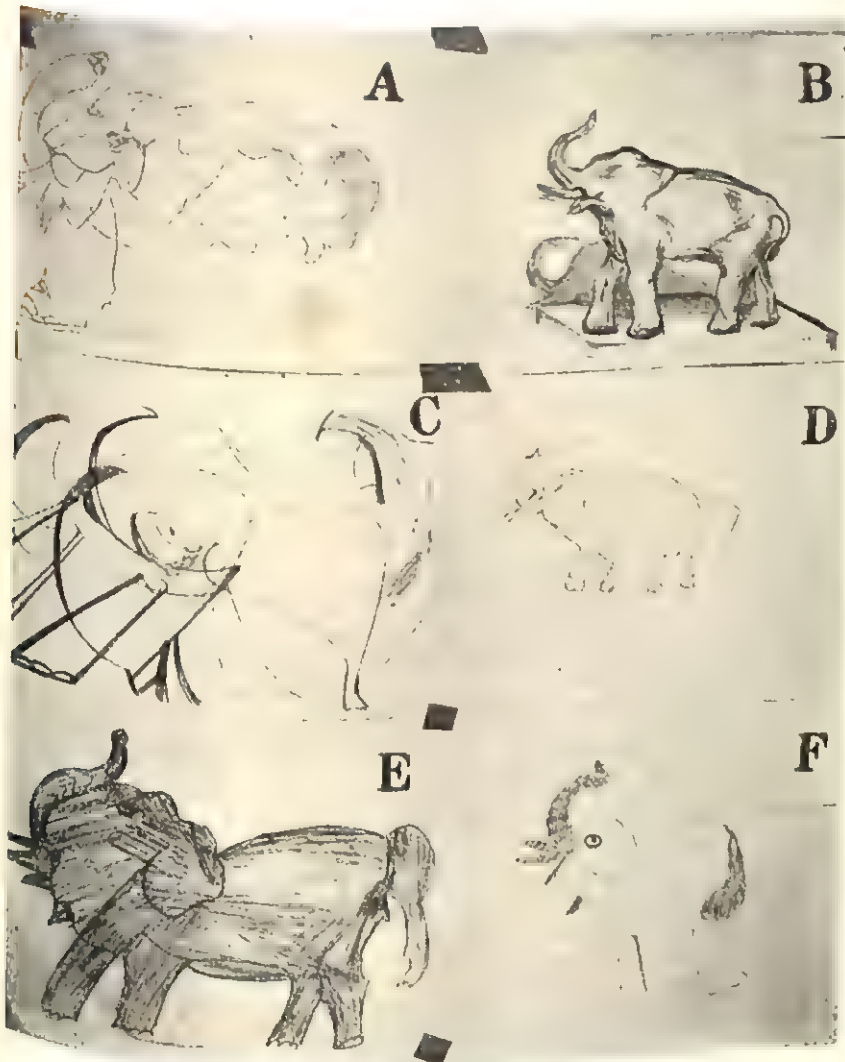


FIGURE 1

FIVE DRAWINGS BY PSYCHOTICS AND ONE BY A NORMAL SUBJECT, REFERRED TO IN THE CASE HISTORIES IN TEXT  
A-4, B-6, C-1, D-3, E-5, F-2.

since age 12; frequent and worsening temper tantrums; ran away and was found with men. Had illegitimate child. Restless, violent, confused at times. Once smashed a glass tumbler because of its "eyes." Frequent releases during which she is able to work until new onset of violence. States of wild, senseless activity; compulsive hand-soaping and wiping. Homicidal and suicidal. EEG looks like petit mal.

*Case 5.* Female, Age 41, 12th grade, *SVA* 15.5, *CQ* 60, on Wechsler, hosp. 131 mo., general paresis. Was healthy child; good school work. Broken home. Extramarital relations with man of own age, who married someone else, then returned to her. Blood test pos. at age 24. Errors in work, violence, homicidal tendencies, inarticulate speech, delusions of grandeur. Heard of project and volunteered for drawing and Shipley Test.

*Case 6.* Female, normal college student, age 21, 2½ years college at time of drawing. Since then highly successful completion of college, successful graduate work, successful professional career and successful marriage. Many conferences with competent psychologists by the senior writer concerning this person reveal very adequate adjustment to life in general.

The writers have no doubt that most readers will have little difficulty in selecting the drawing of *Case 6*. The extent to which the other drawings display clinical signs which make them distinguishable as to the several pathologies is a matter of interest, and the writers would welcome correspondence from colleagues who are engaged in projective work. The question of whether or not discrimination among such drawings is possible when the items selected are not so extreme will be the topic of a later report, after the necessary data have been accumulated.

#### REFERENCES

1. CRANNELL, C. W. The validity of certain measures of art appreciation in relation to a drawing task. *J. of Psychol.*, 1953, **35**, 131-142.
2. SHIPLEY, W. C. Shipley Institute of Living Scale for Measuring Mental Impairment, Manual of Directions. Hartford, Conn.: Institute of Living, 1940.
3. VERNIER, C. M. Projective Test Productions: I. Projective Drawings. New York: Grune & Stratton, 1952.
4. WHITMYRE, J. W. An investigation of the significance of artistic excellence in the judgment of adjustment inferred for human figure drawings. Unpub. Doct. Diss., Univ. Pittsburgh, 1952.

*Department of Psychology*  
*Miami University*  
*Oxford, Ohio*

## TRANSFER OF SKILL ON A FOLLOWING TRACKING TASK AS A FUNCTION OF TASK DIFFICULTY (TARGET SIZE)\*

*Department of Psychology, University of Rochester*

---

RUSSEL F. GREEN<sup>1</sup>

---

### A. INTRODUCTION

A considerable amount of evidence has accumulated to show that there is generally more transfer from a difficult to an easy version of a task than from an easy to a difficult version (1). This result has been obtained in such tasks as problem solving, discrimination learning, and serial motor skills.

There is still, however, reason for questioning whether this will prove to be universally so, although exceptions may depend on our definition of difficulty. Gibbs (1) found that changes in the size of the visual stimulus (length of pointer on a one-dimensional compensatory tracking task) and gear ratio did not effect learning times, rate of learning, or amount of skill transferred. It appeared in Gibbs' study that quadrupling the size of the visual stimulus was not attended by a change in difficulty in the tracking task when difficulty is defined as "time on target" (holding a needle within a space of one degree of arc).

The change in target size introduced changes in both the perceptual and physical requirements of the task (i.e., changed the "sizes" of both the display and the response). Smaller errors should have been more readily perceived when the target was magnified. The operator's hand moved through different distances and at different speeds to effect a correction of errors which would be visually the same for the two targets. Yet skills learned on one size pointer transferred completely to the other size pointer. Even though net proficiency did not change as a function of varying the perceptual and physical requirements of his task, transfer effects could still have been expected to be unequal.

Why the changes of visual stimulus Gibbs introduced did not produce differing degrees of proficiency and of transfer is a moot question. His

---

\*Received in the Editorial Office on December 3, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This study was conducted as part of a research program on human motor performance carried out under Contract No. N6onr-241, T.O. VI between the University of Rochester and the Special Devices Center, Office of Naval Research, to whom these results have been reported in Technical Report No. SDC 241-6-13.



task was a simple one. How fully will his findings be verified when the tasks employed are intrinsically more difficult? It is conclusively known for instance that varying the target size on a following tracking task such as the SAM Two-Hand Coördination Test (2, p. 177 ff.) does produce significant differences in time on target scores. But does varying the target size introduce sufficient differences in the perceptual and physical demands of the task to obtain differences in transfer of skill?

#### B. STATEMENT OF THE PROBLEM

It was our purpose then to investigate the relative amounts of transfer of skill on a following tracking task as a function of task difficulty (target size). The specific hypothesis was that transfer should be greater from a difficult to an easy motor task (Two-Hand Coördination) than it is from an easy to a difficult task. The only characteristic varied in the response-display relationship was that of target size. No change in apparent effect on the target follower as a result of a movement of a given extent of the controls was introduced.

#### C. SUBJECTS

The Ss were 64 male students of the Rochester Institute of Technology.<sup>2</sup> Almost all of them were between 20 and 30 years of age. Ss were assigned to the initial experimental conditions in the order of their appearance. They were paid for their services.

The Ss could be assumed to be relatively naïve as far as psychological tests are concerned but many of them were definitely not naïve, mechanically speaking. Many had had experience with handwheel or crank controls in various situations. It is quite possible that a higher initial level of skill was obtained in this group than in the general population.

#### D. APPARATUS

The apparatus used was a modified SAM Two-Hand Coördination Test (CM101B). A detailed description of the original test may be found in AAF Report No. 4 (2) and details of the modifications may be found in the report of a previous study from this laboratory (4). The S's task was to keep a target follower anywhere on top of a small brass target. The target was fastened to a black turntable which rotated slowly in the horizontal plane in a clockwise direction. The turntable contained a curved slot in which the target moved back and forth in an irregular manner.

<sup>2</sup>We are indebted to Mr. B. G. Culver and staff of the Printing and Advertising Department of the Rochester Institute of Technology for their coöperation in helping to recruit and arrange for testing the subjects.

The rotary movement of the turntable and the irregular movement of the target in the slot combined to give the target an irregular circular course or path during each trial. The course cams are arranged in such a way that the target follows a different course each trial for four trials and then the same four trial courses are repeated.

The *S* moved the target follower by rotating two cranks. The directions of rotation and the planes of rotation of the two cranks used in the present study were those found to be optimum in previous studies (4, 5). The left crank moved the target follower left and right and the right crank moved it toward and away from the subject. The left crank was rotated in a plane parallel to the frontal plane of the body, and the right crank was rotated in the vertical plane parallel to the sagittal plane of the body.

Four round brass targets were used. Their diameters were  $\frac{7}{8}$ -inch,  $\frac{5}{8}$ -inch,  $\frac{3}{8}$ -inch, and  $\frac{1}{8}$ -inch. The  $\frac{5}{8}$ -,  $\frac{3}{8}$ -, and  $\frac{1}{8}$ -inch targets each had a  $\frac{7}{8}$ -inch white collar attached to them. Preliminary observations had indicated that the  $\frac{1}{8}$ -inch target was so small that it was often difficult to tell where the target was in relation to the target follower. The white collar was intended to serve as an additional visual cue to assist the *S* in determining his nearness to the target. White collars were put on the  $\frac{3}{8}$ -inch and  $\frac{5}{8}$ -inch targets to hold constant the visual extent of the target and additional cue areas of all four tasks.

It will be noted that the range of target sizes is quite large. The diameter of the largest target is seven times that of the smallest. The area of the largest target is 49 times the area of the smallest.

Each trial was approximately one minute (actually .985 minutes) in duration and consisted of one complete revolution of the turntable which carried the target. The time on target score was obtained by means of a Standard Electric Timer calibrated in thousandths of a minute. Error scores were obtained by connecting a counter to the clock in such a way that it would count once each time the timer stopped running—that is, each time the *S* permitted the target follower button to lose contact with the target.

#### E. PROCEDURE

A modification of Woodworth's experimental Plan 5 (6) was used to study the transfer effects of interest in this investigation. Briefly, each of four groups was given 16 learning trials on an easy (or hard) target and then was given eight transfer trials on a hard (or easy) target. The transfer design can be symbolized briefly as in Table *d*.

TABLE A

	16 Trials	8 Trials
Group I	Target A ( $\frac{7}{8}$ " )	Target D ( $\frac{1}{8}$ " )
Group II	Target B ( $\frac{5}{8}$ " )	Target D ( $\frac{1}{8}$ " )
Group III	Target C ( $\frac{3}{8}$ " )	Target A ( $\frac{7}{8}$ " )
Group IV	Target D ( $\frac{1}{8}$ " )	Target A ( $\frac{7}{8}$ " )

It was necessary to equate the groups in order to insure equal ability at the outset. Since transfer effects between the targets were unknown, it was decided to have each *S* run one preliminary trial on each of the four targets.

For the preliminary trials it was believed that target sequences and course sequences should be systematically controlled, in order to control target-course interactions. The resulting experimental design was a modified type of Greco-Latin square for the preliminary trials, with four sequences of target sizes and four sequences of target courses. Every target sequence occurred once with every course sequence, and each target occurred four times with each course, once for each trial.

This design holds targets constant for courses and courses constant for targets. It also holds target sequences constant for course sequence and vice versa. It is in effect a cross between a Greco-Latin and a factorial design. The factors of target size, target course, target sequence, course sequence, and learning effect are all orthogonal to or statistically independent of each other. While some interactions are orthogonal to each other and to the main effects others are still confounded with the main effects and with all other interactions as well as with each other. The complete permutation between target size sequences and target course sequences, however, should have balanced some of the confounding effects that might accrue in designs such as this.

The counterbalancing demand made on the trials which were preliminary to the transfer study made it possible to utilize these trials to investigate the possibilities of course differences and target sequence and course sequence differences in difficulty.

The sums of the four preliminary trials were used to equate the groups for the transfer study. Since the sums of scores for the four trials were not necessarily equivalent from one target sequence to another, each group was matched according to target sequence. Thus each of the four groups had equal numbers of individuals with the same target sequence background, but varied somewhat in terms of course sequence and target-course combinations. Even though one of the original matched *Ss* had to be replaced,

the four groups resulting were very closely matched; the mean differences did not exceed .37 seconds per trial.

The procedure for each *S* was the same. Each *S* was tested in two sessions of 15 and 40 minutes separated by approximately two weeks. The time interval between testing sessions varied somewhat from *S* to *S* (ranging from one to three weeks in a few instances).

The instructions to *Ss* were essentially the same as those used by the AAF and reported in full in AAF Report No. 4 (2, p. 192). An additional instruction should be mentioned. The *Ss* who transferred from the two larger targets ( $\frac{7}{8}$ " and  $\frac{5}{8}$ " to the  $\frac{1}{8}$ " target tended to become very discouraged after two or three trials on the  $\frac{1}{8}$ " target. In order to maintain motivation *E* introduced the following statement as a standard procedure when he put the  $\frac{1}{8}$ " target on the turntable:

It has been found in the past that some people become discouraged when they try to follow this target and find that they cannot stay on it as well as they can the larger target. If you just realize that a good performance on this target is very different from a good performance on your other target and keep working you will do all right.

After instituting this remark no obvious instances of giving up were observed. The remark was not instituted until after either five or six *Ss* (out of 32) had been run on the  $\frac{1}{8}$ -inch target. No special remarks were made in connection with any other transfer situation.

## F. RESULTS

The results are divided into two sections. The transfer of skills problem is presented first, with time on target scores and error scores treated separately. Relative difficulty of targets and homogeneity of variance of targets and course difficulty are then discussed in the light of their possible significance to the transfer problem.

### 1. Transfer

a. *Time on target scores.* The primary problem concerned the hypothesis that transfer to a second similar task would be greater when the first task was harder than the second, than it would be when the first task was easier than the second. The means of each of the groups for each trial are presented in Table 1. The means for the preliminary trials, are also included in this table (Table 1A). Each group is represented once in each row and in each column in the table of preliminary trial means. Since individuals were assigned to these groups at random, their mean scores should not differ more than would be expected by chance at any trial.



TABLE 1  
MEAN TIME ON TARGET  
(In milliminutes)

Table 1A Preliminary Trials (N = 16)					Table 1B Learning Trials (N = 16)				
Target trial	7/8"	5/8"	3/8"	1/8"	Target trial	7/8"	5/8"	3/8"	1/8"
a	417	231	155	034	Group	I	II	III	IV
b	554	422	209	059	1	674	584	345	127
c	570	385	323	070	2	709	648	451	131
d	608	531	322	115	3	761	683	459	195
Table 1C Transfer Trials (N = 16)					4	753	635	432	147
Group	I	II	III	IV	5	766	684	504	206
Target trial	1/8"	1/8"	7/8"	7/8"	6	798	714	527	184
1	340	372	908	904	7	854	748	551	259
2	286	330	937	944	8	805	726	497	213
3	317	379	924	945	9	840	790	638	305
4	291	310	925	949	10	895	826	665	269
5	312	360	936	943	11	900	843	650	294
6	296	291	919	941	12	866	814	619	293
7	359	396	931	944	13	881	794	631	317
8	299	327	931	945	14	920	834	691	273
					15	918	867	709	350
					16	907	848	685	296

The trends represented by the means of Table 1 (A, B, and C) can be observed much more readily in graphical form, as presented in Figure 1. A number of things concerning Figure 1 (and Table 1) should be noted. The individuals comprising the several groups are different between the preliminary trials and the learning trials; they are also different between the learning trials and the transfer trials. One-fourth of the individuals in each preliminary group were later assigned to each of the subsequent experimental groups. The groups who tracked the 7/8- and 5/8-inch targets for the 16 trial learning series transferred to the 1/8-inch target for their last 8 trials. Those who tracked the 1/8- and 3/8-inch targets transferred to the 7/8-inch target for their last 8 trials. A two-week interval separated Trial d and Trial 1. A 15-second interval occurred between each of the remaining trials except for a one-minute interval between Trials 8 and 9 and a two-minute interval between the 16th learning trial and the 1st transfer trial.

It will be noticed that in spite of the factors just mentioned, learning appears to progress uniformly throughout the training. In order to illustrate this fact more clearly curves were fitted to the first 20 trials (the four preliminary and the 16 learning trials). These fitted curves were then extrapolated out to the 28th trial. The fitted curves are presented in Fig-

ure 1. The extrapolated portions of the curves are assumed to represent the levels of performance that would have been achieved had the group continued to practice on that target.

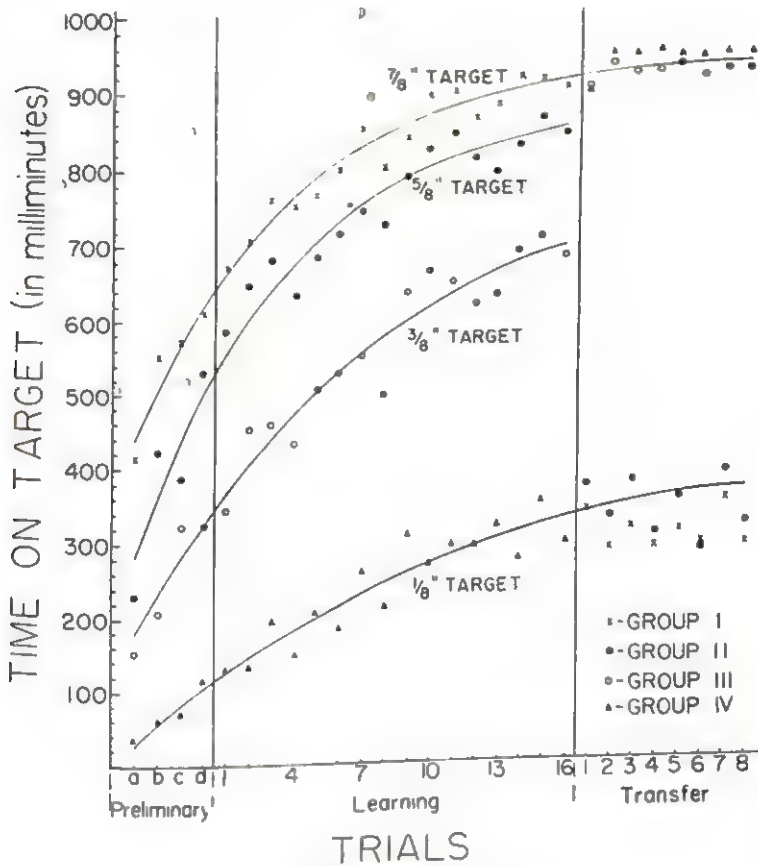


FIGURE 1  
LEARNING CURVES AS TIME ON TARGET

Inspection of these curves shows that transfer of skill from the two smallest targets to the  $\frac{7}{8}$ -inch target was positive and essentially complete. The points representing the performance of the group which had its prior practice on the  $\frac{1}{8}$ -inch target appear slightly higher than those of the group who practiced on the  $\frac{3}{8}$ -inch target. The *t*-ratio of the difference between the combined mean for the eight trials of each of the two groups is however only .74.

Transfer of skill gained from practice on the  $\frac{7}{8}$ -inch target to the  $\frac{1}{8}$ -inch target appears to be somewhat less than that transferred from the  $\frac{5}{8}$ -inch

target to the  $\frac{1}{8}$ -inch target. But again the  $t$ -ratio of the difference between the combined means of the eight trials for the two groups proves to be insignificant (.70). Since the group who had prior practice on the  $\frac{5}{8}$ -inch target appears to have performed at the level which would be predicted from the extrapolation of the fitted curve, it may be assumed that transfer from the easy tasks to the hard task was positive and essentially complete.

It should be noted, however, that in each case, hard to easy and easy to hard, the tendencies to deviate from these conclusions, although not statistically significant, are in both cases in the direction of the hypothesis being tested.

This slight agreement with the original hypothesis may, however, be a result of motivational differences. It will be recalled that special instructions had to be instituted when the  $S$ s transferred from the easy to hard task to keep the  $S$ s from giving up. This tendency to become discouraged did not occur when the  $S$ s transferred from the hard to the easy task.

There are very regular fluctuations in the curves which occur from trial to trial but these are almost certainly largely associated with course difficulty, or interaction effects, or both. For instance, for the  $\frac{1}{8}$ -inch target, scores on Trials 1 and 3 tend to be higher than on Trials 2 and 4 in any given block of four trials following the preliminary trials. For the three larger targets the pattern in each block of four trials is one of rising scores for Trials 1, 2, and 3, and then falling at Trial 4. The reason the course fluctuations do not extend back through the preliminary trials is that due to the experimental design all four courses are represented in the sums for each trial. This rules out differential effects from courses (holds course variability constant for these trials). This design was not carried over into the learning and transfer trials. There only one course is associated with a given trial. (The same target is used across all four courses.)

*b.* Error scores. The mean number of times per trial the  $S$ s lost contact with the targets is presented graphically in Figure 2. The composition of the groups and the intervals between trials are the same as for Table 1 and Figure 1.

Inspection of Figure 2 suggests that the error score functions are complex. All four curves rise initially. The  $\frac{7}{8}$ -inch target curve reaches its peak and starts to diminish very quickly, perhaps between the second and third preliminary trials. The  $\frac{5}{8}$ -inch target curve begins to diminish between the second and third learning trials. The  $\frac{3}{8}$ -inch target curve starts down around the seventh learning trial. The  $\frac{1}{8}$ -inch curve is still increasing at the end of 28 trials.

It is apparent, however, that the error curves show a regular progression and exhibit no gross fluctuations as group composition or as time interval varies. The evidence from the error scores, then, favors the conclusion that transfer of skills from one target size to another is positive and probably complete.

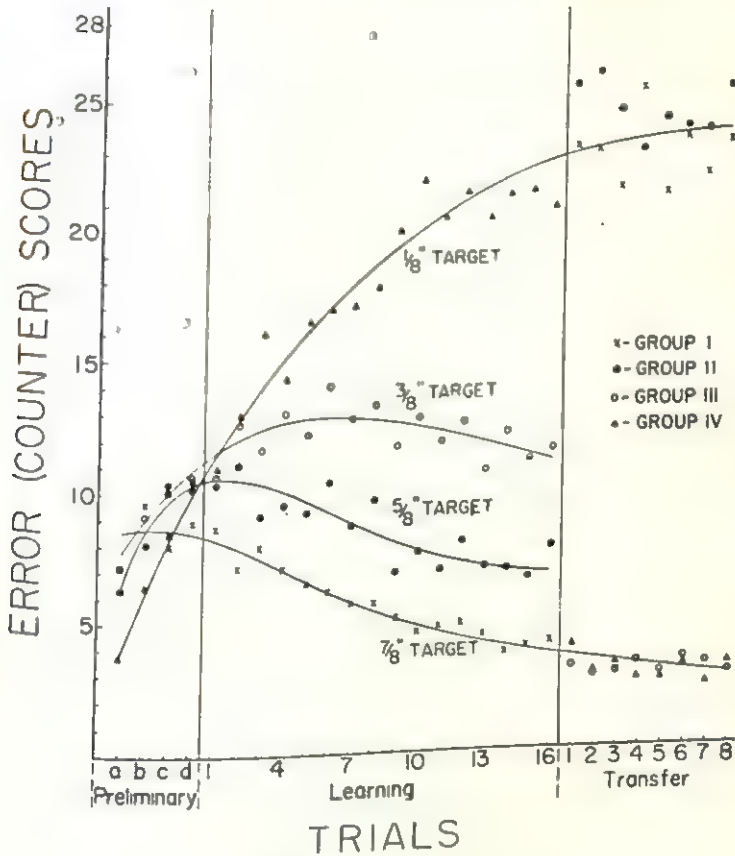


FIGURE 2  
LEARNING CURVES AS ERROR (COUNTER) SCORES

## 2. Relative Difficulty of Targets and Homogeneity of Variance of Target Scores

a. *Time on target scores.* The mean time on targets and variances of scores on each target, for the preliminary trials only, are presented in Table 2. The homogeneity of the variance of the preliminary scores was tested by Bartlett's Method. The corrected Chi-Square was 78.932, which for three degrees of freedom was highly significant. Thus the variances of the pre-



liminary scores obtained on the four targets almost certainly did not come from the same population of scores. When the data from the smallest target were excluded the variances of the scores obtained on the three larger targets were found to be homogeneous (Chi-Square being 2.65 which for two degrees of freedom is not significant).

TABLE 2  
MEAN TIMES ON TARGETS AND VARIANCES OF SCORES  
(For the four preliminary trials.)

Target	Mean	Variance
$\frac{7}{8}$ -inch	537	28,034
$\frac{5}{8}$ -inch	392	30,856
$\frac{3}{4}$ -inch	252	20,689
$\frac{1}{8}$ -inch	70	2,856

Since the  $\frac{1}{8}$ -inch target appeared in all four sequences and in each position of each sequence, it was decided to analyze the variance of the scores of the four preliminary trials twice, once with the  $\frac{1}{8}$ -inch target score variance included, and once with it excluded. The two analyses yielded the same pattern of significant and non-significant  $F$ 's throughout, with one minor exception, thus inclusion of the  $\frac{1}{8}$ -inch target apparently did not affect the outcome of the analysis to any important extent in spite of the fact that its variance was significantly different from that of the other three targets. The more conservative analysis (which excludes the  $\frac{1}{8}$ -inch target score variance) is presented in Table 3.

TABLE 3  
ANALYSIS OF VARIANCE OF THE PRELIMINARY TRIALS  
( $\frac{1}{8}$ " target scores excluded.)

Source	df	MS	F
1. Target size	2	1,301,027	153.9**
2. Target sequence	3	13,521	N. Sig.
3. Course	3	4,769	N. Sig.
4. Course sequence	3	40,857	4.83***1
5. Ordinal position (Practice effects of target size and course)	3	390,235	46.15**
6. Individual differences	45	47,115	5.574**
7. Square unq. for target	3	64,460	7.625**
8. Square unq. for course	6	22,395	2.649* (about 2%)
9. Error	123	8,454	
10. Total	191		

\* = Significant between 5 per cent and 1 per cent levels.

\*\* = Significant at or beyond the 1 per cent level.

<sup>1</sup>Course sequence variance is not significantly different from individual differences MS ( $F = 0.867$ ).

The tremendous *F*-ratio obtained makes it obvious that initially the targets used constituted a wide range of difficulty even though the target size variance may include variance from any of several interactions plus an error component. The smallest *t*-ratio between target means is approximately 8.

The significance of the square uniqueness term for targets indicates that there were significant interactions between targets and one or more other variables. It is apparent from the analysis that the target-by-course and target-by-target sequence interactions probably are not significant. This follows since the course and target sequence mean squares are not significantly different from the error term. Of course, if the error term were spuriously high as a result of containing interaction variance, this conclusion could be reversed. But target-by-course and target-by-target sequence interaction variances could not be higher than the target size sequence and course variances except by chance.

It seems more likely that there are significant target-by-course sequence and (or) target-by-subject interactions. The design of this experiment makes it possible to compute various interaction terms including target-by-course sequence. These interaction terms would be of doubtful interpretive value, however, since they are all confounded with trial-by-target sequence, trial-by-course sequence, and various interactions involving subjects.

It has been shown that initially the targets used constituted a wide range of difficulty. It is also true that the magnitude and significance of the differences between the means and variances of these four target sizes shifts as practice proceeds. These shifts are further influenced by the artificial ceiling of the scores obtainable. Consider, for instance, the means and variances of scores on the various targets on the sixteenth learning trial shown in Table 4.

TABLE 4  
MEAN TIMES ON TARGETS AND VARIANCES OF SCORES  
(For the 16th learning trial.)

Target	Mean	Variance
$\frac{7}{8}$ -inch	907	8,054
$\frac{5}{8}$ -inch	848	5,640
$\frac{3}{8}$ -inch	686	36,729
$\frac{1}{8}$ -inch	296	13,211

The difference between the  $\frac{7}{8}$ -inch and  $\frac{5}{8}$ -inch targets falls just short of the 5 per cent level of significance ( $t = 1.954$ ). All other differences between means are significant. The difference between the  $\frac{7}{8}$ - and  $\frac{1}{8}$ -inch targets is numerically greater than it was early in practice. The variance

of the  $\frac{3}{8}$ -inch target is very significantly larger than that of any of the other three targets, while those of the other three are homogeneous in this respect. These shifts of apparent difficulty of the targets is noteworthy in that, in spite of them, transfer of skills from one target to another appeared to be virtually complete. This would imply that even though the subjects who practiced on the  $\frac{7}{8}$ -inch target approached the artificial ceiling of efficiency imposed by the scoring method used, their actual efficiency was continuing to improve at the same rate it would have if there had been no such maximum possible performance (defined as time on target during a trial).

*b. Error scores.* The results of an analysis of variance of the counter (error) scores obtained on the four preliminary trials were essentially the same as the time on target results. The major difference was that there was less evidence of significant interactions between variables. The square uniqueness estimates were not significant.

Even though the subjects made significantly varying numbers of errors as target size changed, these differences did not parallel those from time on target. The average number of errors (during the preliminary trials) for each of the four targets from the largest target to the smallest were 8.98, 9.98, 9.86, and 7.70 respectively. Errors occurred a very significantly greater number of times on the  $\frac{5}{8}$ -inch and  $\frac{3}{8}$ -inch targets than on the  $\frac{1}{8}$ -inch target ( $t = 4.61$  and  $4.36$ ). Errors on the  $\frac{7}{8}$ -inch target occurred significantly less often than they did on the  $\frac{5}{8}$ -inch target ( $t = 2.01$ ), and more often than they did on the  $\frac{1}{8}$ -inch target ( $t = 2.59$ ).

The subjects stayed on the three larger targets far more than they did on the  $\frac{1}{8}$ -inch target yet they made more "errors" on the larger targets during the four preliminary trials. The difficulty here is that it was harder for the subjects to get on the  $\frac{1}{8}$ -inch target so they could make an error than it was with the three larger targets.

*c. Course differences.* Analysis of the variance of the four preliminary trials indicated no significant differences between difficulty of trial courses as measured by mean time on target. The mean times on targets for each of the four courses, with target size held constant, was 319, 310, 321, and 301 milliminutes respectively.

No analysis of relative course difficulty throughout the learning trials was possible. The curves presented in Figure 1 suggest that such differences may exist at later stages of practice even though they were not manifested in the first four trials.

Analysis also showed that errors (i.e., number of times the subjects lost contact with the target) were no more frequent on one course than they

were on any other course during the four preliminary trials. It seems probable that varying course difficulty is not an important factor during the first few trials on the SAM Two-Hand Coordination Test but that these differences may become more important as learning progresses.

## G. DISCUSSION AND RECOMMENDATIONS

### 1. *Transfer*

*a. Time on target scores.* The results indicate that practice on a target of one size on this tracking task is essentially equivalent to practice on a target of any other size used in this experiment. It seems likely that transfer of training on one target is positive and essentially complete to another target of different size when level of training is measured either as time on target or as number of times contact with the target was broken.

It is true that the design of this experiment does not completely justify the above conclusion. Considerable corroboration has come, however, from an Air Force study, recently reported in abstract form (3). A more completely balanced design was used in the Lackland study although they used only two target sizes ( $\frac{7}{8}$ -inch and  $\frac{3}{8}$ -inch). The results from the two studies are in essential agreement. The present study extends the generality of the conclusion over a wider range of target size values. The results of these two studies are also in agreement with Gibbs' results (1). Changes in the size of the visual stimulus through a quite large range with its concomitant change in the precision of adjustive movements required, apparently do not affect rate of learning or amount of skill transferred.

Even though the mean time on target scores are very significantly different and so can be considered discrete tasks from the statistical standpoint, it can be held that from the standpoint of what is learned, all four tasks are the same in the sense that for practical purposes they develop the same skills. It would seem reasonable to recommend, if the size of the target to be tracked is the prime variable by which the level of difficulty of the training task is to be controlled, that for a wide range of target sizes the initial training can be on a large (easy) target. Targets can be made smaller (harder) as training progresses. Learning will be equally fast. If the training sessions are equally long motivation should be enhanced with larger targets since the demands on the trainee will be lower and apparent success will be greater.

*b. Error scores.* The counter scores must be considered essentially error scores since the subjects' task was to keep the target follower button on the target and so minimize the counter score. The counter score gives a measure



of subjects' failures to keep on target. The complexity of the error functions suggests, however, that there are really two tasks involved in tracking tasks such as that on the SAM Two-Hand Coördinator. One task is to get back on the target as soon as possible after losing contact with it. The second task is to stay on the target after having got on it. It is quite likely that these two tasks are not only different but that they also will change in relative importance as a function of learning and as a function of task difficulty.

It is true that superimposed upon the task of correcting the observed error is the task of following the target through a portion of its course while the error is being diminished. The correction of the gross error may be the most important element in the task when such an error exists. When the subject is on target the most important element in the task may be that of attempting to "anticipate" where the target is going next and to prevent the occurrence of errors (loss of contact). If the anticipation of future events may be considered an essential element in tracking, getting on target may involve correcting for errors to a greater extent than it involves tracking while staying on target may involve tracking to a greater extent than correcting for errors—although an error cannot be corrected without concomitant tracking (anticipation of future target course).

The above analysis seems to apply fairly well to the three larger targets used for this study. But the smallest target ( $\frac{1}{8}$ -inch) presents a special problem. It is possible that the counter score for the  $\frac{1}{8}$ -inch target could be considered either a "number of hits" score, and so a "correct" score, or an error score. It might better be considered a correct score since the target is so small that it is very difficult to keep the follower on the target once it has been placed there. The task the subject is actually performing most of the time, then, is one of trying to hit, to get on, the target (that is, correct for error).

Here we have a clue to what seems to be a basic qualitative difference between the task of tracking the  $\frac{1}{8}$ -inch target as compared to the task of tracking each of the other targets—a difference which has impressed the experimenter in observing subjects perform these tasks. With each of the other targets the task quickly becomes one of "trying to stay on target" while the task presented by the  $\frac{1}{8}$ -inch target remains, primarily, one of "trying to get on target."

Some support for this hypothesis can be drawn from the two families of curves. It has been noted that the counter scores show an initial rise for all four targets. For each target size the most important portion of the sub-

jects' task initially is to get on the target. Learning is very rapid with the larger targets so that the most important element of the task quickly changes to one of staying on target—with getting back on target becoming a comparatively less important problem. It will be noted that the peak of each of the three curves of the counter scores comes at a trial on which the time on target is at or only a little more than one-half the total possible time. Apparently the relative importance of the two tasks changes when the subject achieves approximately 50 per cent efficiency (when efficiency is defined as time on target).

At this point, one may well ask how it is that transfer between the  $\frac{1}{8}$ -inch target and the  $\frac{7}{8}$ -inch target is positive and essentially complete. If there is an important difference between correcting for errors and the element of anticipation (staying on target), something less than complete positive transfer might be expected. Both tasks are doubtless being practiced throughout much of each trial regardless of target size. An error can be made and corrected without losing contact with a large target. Such error cannot be corrected, however, without including a prediction as to where the target is going to be in the near future. One would think that these two tasks would not be practiced to the same extent or in the same proportion on each of two targets of markedly different size. An explanation might lay in the fact that it is probably true that varying weights of the same abilities can produce the same or similar performances. Of course, it is also possible that such tendencies for the transfer of skills to differ as were observable were due to this apparent qualitative difference between the tasks. But these tendencies were slight and insignificant.

#### H. SUMMARY

The hypothesis tested was that transfer should be greater from a difficult to an easy two-hand coördination task than it is from an easy to a difficult task. Sixty-four young adult males served as subjects. The apparatus used was a modified SAM Two-Hand Coördination Test.

Each *S* was given four preliminary trials, one on each of four targets ( $\frac{7}{8}$ -inch,  $\frac{5}{8}$ -inch,  $\frac{3}{8}$ -inch, and  $\frac{1}{8}$ -inch in diameter). These scores were used to divide the subjects into four groups equated for initial ability. The transfer design used is shown in Table *A*.

Transfer of training both from large (easy) targets to small (difficult) targets and from small to large targets was positive and essentially complete. This observation held both for a time on target score and for an error score (defined as number of times contact with the target was broken).

The results indicate that the hypothesis should be rejected for this task. Practice on a target of one size is essentially equivalent to practice on a target of any other size used in this experiment. However, it seemed possible that the limits of the range of difficulty across which this result will be obtained (if there are such limits) were very nearly included in this experiment. So, even though the mean time on target scores are very significantly different and thus can be considered discrete tasks from the statistical standpoint, it can be held that from the standpoint of what is learned all four tasks are the same in the sense that for practical purposes practice on them apparently develops the same or equivalent skills.

## REFERENCES

1. GIBBS, C. B. Transfer of training and skill assumptions in tracking tasks. *Quart. J. Exp. Psychol.*, 1951, **3**, 99-110.
2. MELTON, A. W. Apparatus Tests, *Report No. 4, U. S. Army Air Forces Aviation Psychology Reports*. Washington: U. S. Government Printing Office, 1947.
3. MORIN, R. E. Transfer of training between motor tasks varying in precision of movement required to score. *Amer. Psychol.*, 1951, **6**, 390.
4. NORRIS, E. B., & SPRAGG, S. D. S. Performance on a following tracking task (modified SAM Two-Hand Coördination Test) as a function of the planes of operation of the controls. *J. of Psychol.*, 1953, **35**, 107-117.
5. ———. Performance on a following tracking task (Modified SAM Two-Hand Coördination Test) as a function of the relations between direction of rotation of controls and direction of movement of display. *J. of Psychol.*, 1953, **35**, 119-129.
6. WOODWORTH, R. S. *Experimental Psychology*. New York: Holt, 1938. (P. 180).

*Department of Psychology*  
*University of Rochester*  
*Rochester 3, New York*

# INTENSIVE EQUIVALENCES FOR SUCROSE AND NaCl SOLUTIONS\*

Department of Psychology, Harvard University

J. G. BEEBE-CENTER, M. S. ROGERS, AND W. H. ATKINSON

A preliminary graphical specification of intensive equivalences for sucrose and NaCl solutions was published in 1948 by Beebe-Center and Waddell (1). Figure 1 provides a more thorough specification of this relation.

Squares and triangles represent specific equivalences for four individual subjects obtained by the method of constant stimuli by Beebe-Center and Waddell. The squares refer to experiments in which the standard was a

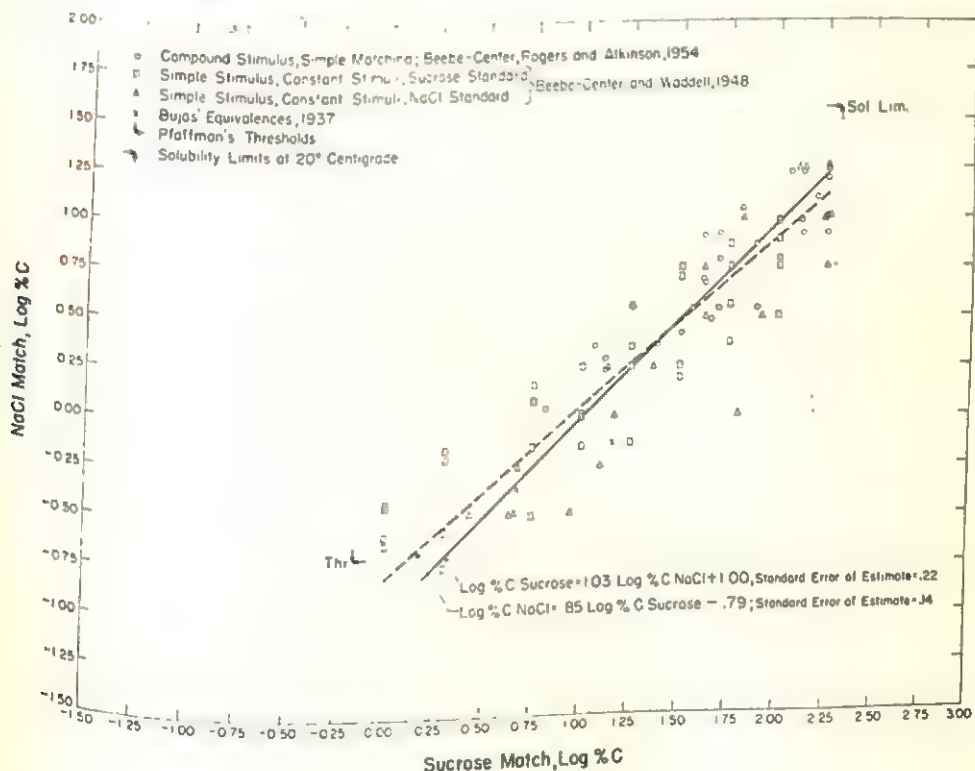


FIGURE 1

(The second member of the equation should start with .85, not 85.)

\*Received in the Editorial Office on December 10, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.



sucrose solution and the variables were NaCl solutions; the triangles, to experiments in which the standard was an NaCl solution and the variables were sucrose solutions. Circles represent specific equivalences for three individual subjects (the authors of this article) obtained by matching compound solutions of sucrose and NaCl for general intensity against a series of simple sucrose solutions and a series of simple NaCl solutions. The two crosses represent intensive equivalences reported in 1937 by Bujas (2).

The arrows labelled "Thr." in the left lower part of the figure represent the sucrose and NaCl threshold values given by Pfaffmann (4). The arrows in the right upper part of the figure, labelled "Sol. Lim.," represent the aqueous solubility limits at 20° Centigrade for sucrose and NaCl (3).

The equations in the figure represent, for the squares, triangles, and circles, the regression of log per cent concentration of sucrose on log per cent concentration of NaCl, and vice versa.<sup>1</sup> Each equation is followed by its standard error of estimate.

A good approximation to the two equations is provided by the statement that, with respect to general subjective intensity, the effectiveness of saline solutions (in per cent concentration) is 10 times that of NaCl solutions (in per cent concentration).

#### REFERENCES

1. BEEBE-CENTER, J. G., & WADDELL, D. A general psychological scale of taste. *J. of Psychol.*, 1948, **26**, 517-524.
2. BUJAS, Z. La mesure de la sensibilité différentielle dans le domaine gustatif. *Acta Inst. Psychol. Univ. Zagreb.*, 1937, **2**, No. 1. Pp. 13.
3. HODGMAN, C. D. (Ed.). *Handbook of Chemistry and Physics*. (18th edit.) Cleveland: Chemical Rubber Publ. Co., 1933.
4. PFAFFMANN, C. Taste and Smell. In S. S. Stevens (Ed.), *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 1143-1171.)

*Department of Psychology*  
*Harvard University*  
*Memorial Hall*  
*Cambridge 38, Massachusetts*

---

<sup>1</sup>Per cent concentration (% C) equals grams solute per 100 cc tap water.

## LYSERGIC ACID DIETHYLAMIDE (LSD-25): IV. EFFECT ON ATTENTION AND CONCENTRATION\*<sup>1</sup>

*Departments of Medicine, Neurology, and Psychiatry, The Mount Sinai Hospital  
New York City*

M. E. JARVIK, H. A. ABRAMSON, AND M. W. HIRSCH

### A. INTRODUCTION

Following Hofmann's discovery of lysergic acid diethylamide (LSD-25) in 1943 numerous investigators attempted to evaluate the effects of this drug. Among these were Busch and Johnson (2), Condrau (3), De Shon (4), Forrer and Goldner (5), Gastaut (7), Rinkel *et al.* (9), Savage (10), and Stoll (13). Physiological, perceptual, and psychological changes were observed clinically and occasionally measured objectively.

Only one study, that of Gastaut (7), bases the conclusion that attention was impaired on specific psychological tests, viz., Cattell, Lahey, and Rorschach. The other findings cited are based only on clinical observations.

The purpose of this investigation, then, is to determine the effects of LSD-25 on indices of the central processes of attention and concentration (8). Speed of movement is another factor that enters into this test, but in subsequent papers to be published in this series we discuss the effect of LSD-25 on speed of movement *per se*.

As a more objective measure of what may roughly be called concentration, cancellation tests have been found useful. The use of these tests has been reported by Thorndike and Woodworth in 1901 (14), and by Sterzinger in 1924 (11), as well as by others. A cancellation test is one in which the subject underlines, crosses out, or in some other way indicates his recognition of certain recurring items or classes of items on a page of repetitive, though randomized material. His accuracy and speed, indices of some rather complex unspecified mental processes, are used as the dependent variables.

The reliability of these tests has been summarized by Whipple (15) who reports reliability co-efficients ranging from 0.60 to 0.97 for the various forms and lengths of cancellation tests. Despite their reliability, however, speed of cancellations has been found to increase with practice according to some investigators (11).

\*Received in the Editorial Office on December 25, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was aided by a grant from the Geschichter Fund for Medical Research. We are indebted to Sandoz Pharmaceuticals, Inc., for supplies of LSD-25 and other compounds.

## B. METHOD

### 1. *Subjects*

The subjects in this study consisted of six male and six female, adult paid volunteers who were considered non-psychotic on the basis of psychiatric interviews and a battery of clinical psychological tests as described in the first paper of this series (1). Subjects who were deemed to have certain psychotic traits were rejected from the group of subjects. Intelligence of the subjects, as measured by the Wechsler-Bellevue Intelligence Scale, ranged from average to superior intelligence for eight of the subjects. The median age was 27, ranging from 21 to 33. Median weight was 135 pounds, ranging from 103 to 175 pounds. Seven subjects were graduate students in psychology or allied fields, two were housewives, one an auditor, one a hair-dresser, and one a public relations worker.

### 2. *Tests*

The cancellation tests used were obtained as printed tests from C. H. Stoelting Company, Chicago, Illinois (12). The following were used: (a) No. 27008-Cancellation Test from Whipple's Manual (letters), (b) No. 27002—The Munsterburg Attention Test (numbers), (c) No. 27001—The Munsterburg Attention Test (letters), (d) No. 27010, from Whipple's Manual (geometric figures), (e) No. 27011, from Whipple's Manual (numbers) and, (f) No. 27009, from Whipple's Manual (letters).

### 3. *Procedure*

*a. Administration of drug.* The drug was administered orally in 75 cc of water at approximately 9:30 A.M. The subjects had had no food since the previous evening. The drug administration took place in an air-conditioned room, in which the subjects were tested in groups of 3, 4, or 5. All the subjects had previously been in this room, and had met the experimenters when they were given preliminary interviews and the clinical psychological tests (1). The subjects sometimes knew the other members of the group. Two experimenters administered the tests.

The tests were administered at four separate testing sessions at least one week apart. The first time the subjects were tested, they received a zero dose of the drug, a placebo consisting of water. Since the drug has no taste, color, or odor, the subjects had no knowledge of the amount of the drug they received. The second and third times the subjects received 50 and 100 micrograms of LSD-25, respectively. In order to measure the effect of re-

peated testing the subjects were given as a final dose zero micrograms of LSD-25.

A light breakfast was given to the subjects one-half hour after the drug was administered. No stimulants such as coffee or tea were permitted, and no smoking was allowed during the experimental day.

*b. Administration of tests.* Between  $1\frac{1}{4}$  to  $1\frac{1}{2}$  hours after ingestion of the drug, the cancellation tests were administered to the subjects. Previous to these tests the subject had spent about 15 minutes taking other tests measuring learning and memory. These results will be described in subsequent papers. The tests were administered in the order listed with the following types of cancellations or underlinings: "a's," "odd numbers," all letters from "a" through "l," "consecutive numbers whose sum was 9," "consecutive letters which made words," "squares," "odd numbers immediately preceded by odd numbers," and "vowels." The subjects were given standard instructions each time and the same form of the tests was used at each of the four testing sessions.

*c. Analysis of data.* The score of each test was the number of correct responses. The scores at zero and 50 micrograms of LSD-25, zero and 100 micrograms, and 50 and 100 micrograms of the drug were then compared by the Wilcoxon non-parametric test of paired replicates (16), to determine whether the score differences between doses were significant.

In order to evaluate the effects of practice with a given test, the scores obtained under the first zero testing and the second zero testing were compared by the same technique. The Spearman rank-order correlation (6, p. 355) was used to estimate the predictability of performance under one dose from the rank under another dose of the drug or placebo.

### C. RESULTS

The tests were divided arbitrarily into two main types: those requiring simple discrimination and those requiring complex discrimination. In the first category the decision to cross out could be made by merely looking at a single letter, number, or form, whereas in making the complex discrimination subjects had to look at two or more of the letters or words before they could decide whether to cross out or underline.

In Figures 1 and 2 the mean scores of 12 subjects on three test days (pre-drug zero, 50, and 100 micrograms) are illustrated. Most tests in Figure 1, but none in Figure 2, show a progressive drop in mean score under increasing drug dose. The simple discrimination tests are plotted in Figure 1, the complex in Figure 2.



Table 1 lists the significance of the change between scores obtained under the first zero dosage and the subsequent two testings with LSD-25, and between the two testings with LSD-25. It can be seen that there is no signifi-

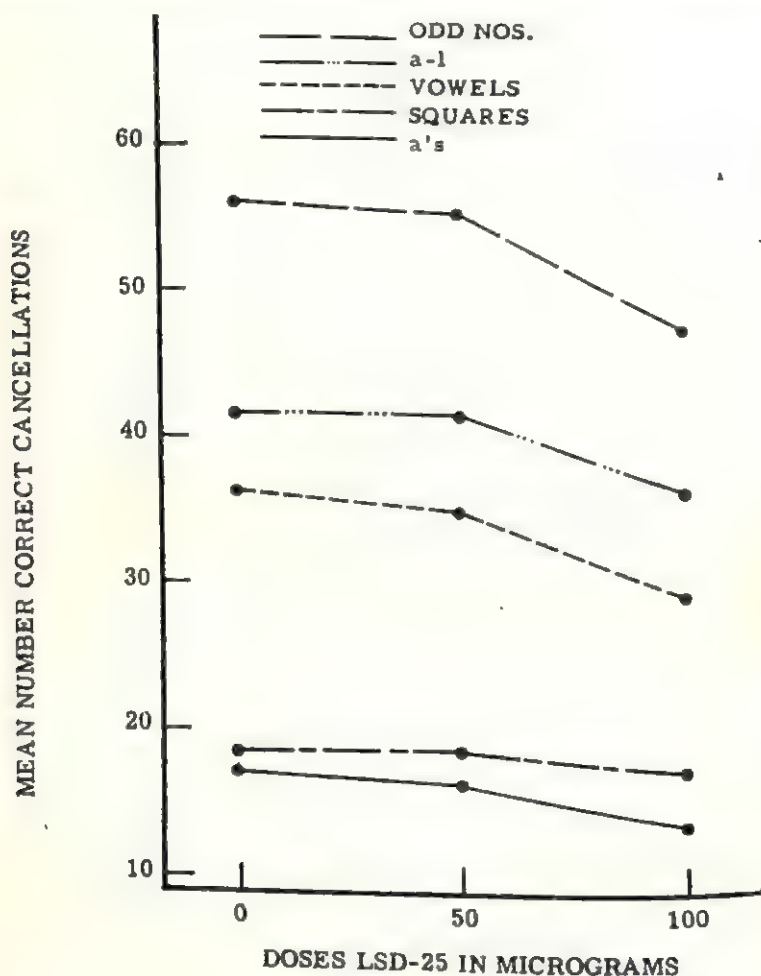


FIGURE 1  
MEAN SCORES ON FIVE TESTS OF CANCELLATION INVOLVING SIMPLE DISCRIMINATION:  
SUBJECTS WERE TESTED UNDER THREE DOSES OF LSD-25

cant difference between the zero and 50 microgram scores on any simple discrimination test, but there is a significant difference between the zero and 100 microgram tests for cancellation of "a's" and "vowels." There is also a significant difference between the 50 and 100 microgram tests for cancellation of "a's," "vowels," and "odd numbers."

For complex discriminations, Table 1 shows that the "word" test scores are significantly different between zero and 50 and between zero and 100 microgram comparisons. The "odd numbers preceded by odd numbers" score

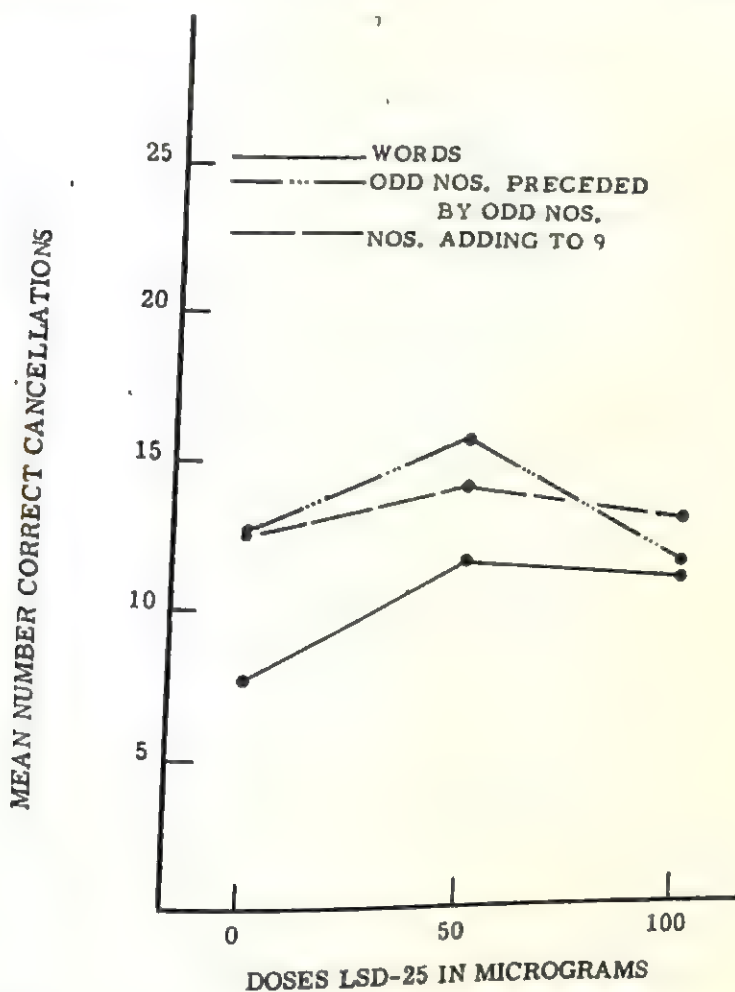


FIGURE 2  
MEAN SCORES ON THREE TESTS OF CANCELLATION INVOLVING COMPLEX DISCRIMINATION:  
SUBJECTS WERE TESTED UNDER THREE DOSES OF LSD-25

is significantly different only for the zero to 50 microgram comparison. These changes were in the direction of improvement under the higher dose. "Numbers adding to 9" scores show no significant changes under LSD-25. Table 2 attempts to show the effects of simple repetition of the tests, divorced from LSD-25 action. When the post-drug zero scores are compared

with the pre-drug zero scores the results illustrated in Figures 3 and 4 show significant differences for all tests except cancellation of "a's" and "squares."

Table 3 shows how constant the relative standing of the subjects was on a given test but under different doses of LSD-25. Four of the eight tests showed significant rank-order correlations between zero and 50 micrograms,

TABLE 1

PROBABILITY OF CHANCE OCCURRENCE OF SCORE DIFFERENCE BETWEEN DRUG DOSES  
(Pre-drug zero, 50 and 100 micrograms LSD-25 compared;  $N = 12$ )

Test	Micrograms LSD-25 compared					
	Zero versus 50 Direction of change	P	Zero versus 100 Direction of change	P	50 versus 100 Direction of change	P
<i>Simple Discrimination</i>						
"a's"	L*	—	L	<.02	L	<.05
"squares"	H**	—	L	—	L	—
"vowels"	L	—	L	.05	L	<.02
"odd nos."	L	—	L	—	L	.05
"a-1"	L	—	L	—	L	—
<i>Complex Discrimination</i>						
"words"	H	<.01	H	<.05	L	—
"odd nos. preceded by odd nos."	H	.05	L	—	L	—
"nos. adding to 9"***	H	—	H	—	L	—

P indicates probability that difference would occur by chance.

— indicates that the probability of the difference occurring by chance is greater than .05.

\*L means lower score under higher dose.

\*\*H means higher score under higher dose.

\*\*\* $N=11$ .

TABLE 2

PROBABILITY OF CHANCE OCCURRENCE OF CHANGE IN SCORES FROM PRE-LSD-25 TO  
POST-LSD-25 TESTINGS ( $N = 8$ )

Test	Direction of change	P
<i>Simple discrimination</i>		
"a's"	H*	>.05
"squares"	H	>.05
"vowels"	H	<.02
"odd nos."	H	.01
"a-1"	H	.01
<i>Complex discrimination</i>		
"words"	H	.01
"odd nos. preceded by odd nos."	H	.01
"nos. adding to 9" **	H	.02

\*H means higher score under Post-LSD-25.

\*\* $N = 7$ .

three were significant between zero and 100, and three were significant between 50 and 100 micrograms.

TABLE 3  
SPEARMAN RANK-ORDER CORRELATION COEFFICIENTS  
(Three doses of LSD-25, related two at a time;  $N = 12$ )

Test	Micrograms LSD-25 compared					
	0 versus 50		0 versus 100		50 versus 100	
	$\rho$	P	$\rho$	P	$\rho$	P
<i>Simple discrimination</i>						
"a's" "	.51	—	.76	.01	.70	.05
"squares"	.63	.05	.68	.05	.87	.01
"vowels"	.67	.05	.35	—	.73	.01
"odd nos."	.35	—	.39	—	.56	—
"a-1"	.89	.01	.50	—	.36	—
<i>Complex discrimination</i>						
"words"	.60	.05	.54	—	.51	—
"odd nos. preceded by odd nos."	.39	—	.64	.05	.38	—
"nos. adding to 9" *	.58	—	.48	—	.44	—

$\rho$  refers to correlation coefficient.

— indicates correlations that could occur by chance more than five times in 100.

\*  $N = 11$ .

#### D. DISCUSSION

In evaluating the effect of lysergic acid diethylamide on the performance on the series of eight cancellation tests, it is important to remember that the same form of each test was used in the four test situations, and that the same experimenters administered the tests in the same way each time.

If no drug had been given during the second and third test days, it is likely that the scores would have become successively higher due to repetition. To assess the effects of practice alone, a large sample of control subjects would be necessary. It is unreasonable, however, to expect a decrease in scores on successive trials. Consequently we must ascribe the impaired performance to the effect of the drug. In fact, if the effects of practice could have been removed, the drug effects might have been even more marked.

All of the complex discrimination tests showed higher scores under 50 micrograms than on the pre-LSD-25 testing (see Figure 2). Although it is possible that the subjects improved in attention and speed under the drug, it is more likely that the increase in scores represents practice. Under 100 micrograms there was a relative decrease.

It is interesting to note that the two tests which did not show a significant practice effect were those which required the least amount of attention and the simplest kind of discrimination—cancellation of "a's" and of "squares." In



both these tests, subjects had only one item to cancel, whereas in the other simple discrimination tests, there were from five to 12 items which could be cancelled.

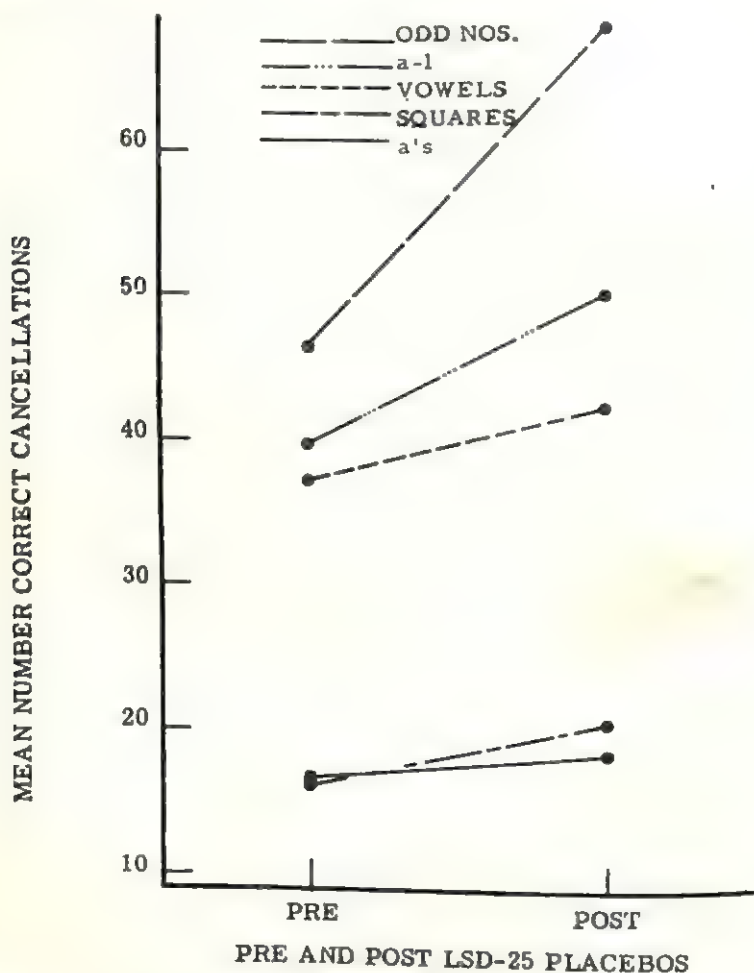


FIGURE 3  
COMPARISON OF MEAN SCORES UNDER PRE-LSD-25 AND POST-LSD-25 PERFORMANCE  
ON FIVE SIMPLE DISCRIMINATION CANCELLATION TESTS

In Table 3 the rank-order correlations for each test between different dosages is given. This shows how consistently the subjects ordered themselves on the same test under different dosages of LSD-25. In a sense, it is a measure of reliability of the tests. All of the correlations are positive and many were significant and fairly high. This suggests that one could predict a

person's rank on a test under one dose of LSD-25 from his rank on the same test under another dose of LSD-25. It is interesting that only some of the tests permit this prediction.

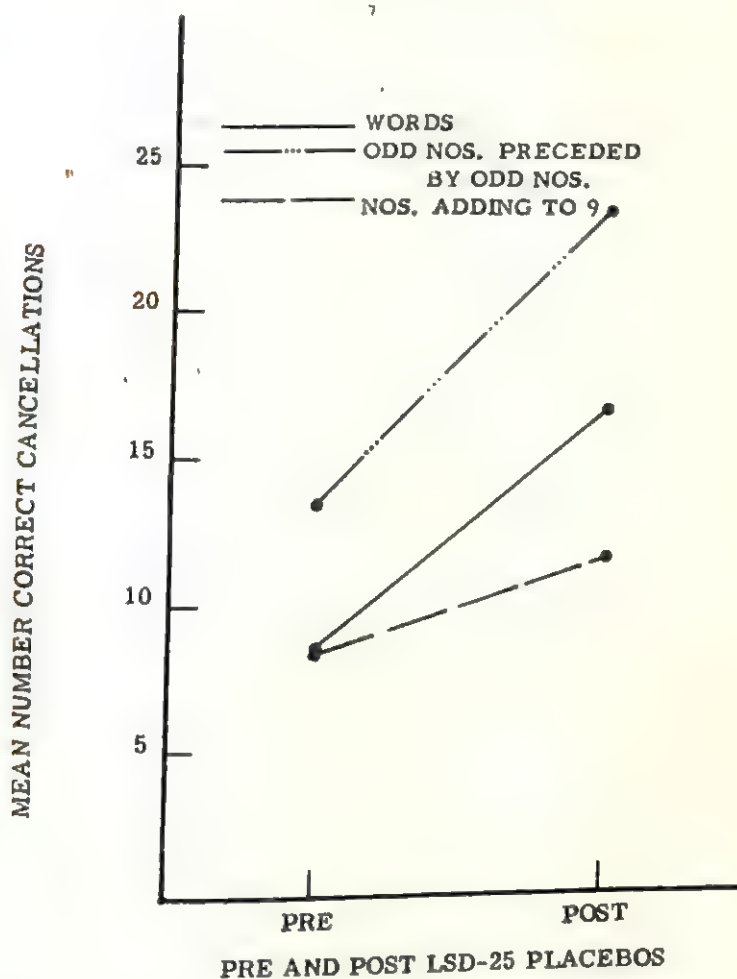


FIGURE 4  
COMPARISON OF MEAN SCORES UNDER PRE-LSD-25 AND POST-LSD-25 PERFORMANCE  
ON THREE COMPLEX DISCRIMINATION CANCELLATION TESTS

### E. SUMMARY AND CONCLUSIONS

On the basis of 12 non-psychotic adults who were tested under a 50 microgram dose of lysergic acid diethylamide, under a 100 microgram dose, and under a zero dose both before and after the LSD-25 tests, with a series of eight cancellation tests, the following conclusions can be made:

1. With the exception of cancellation of "a's" and of "squares" there was a significant increase between pre-LSD-25 and post-LSD-25 zero dose scores for all the cancellation tests.

2. Six out of eight tests showed no significant change when scores of the pre-LSD-25 zero test were compared with the 50 microgram scores. Four of these showed increases and two showed decreases. Two complex discrimination tests showed significant improvement. These were "words" and "odd numbers preceded by odd numbers."

3. Two tests, cancellation of "a's" and "vowels" showed significantly impaired performance when zero and 100 micrograms of LSD-25 were compared. Cancellation of "words" showed a significant improvement in performance under 100 micrograms of LSD-25. Of the remaining five tests, performance on four deteriorated and on one improved, on the average. These changes were attributed to chance.

4. The score decreases under the drug probably would have been even more significant if it were possible to factor out the effects of practice.

5. Performance under a zero dose permits prediction of ranks under 50 micrograms of LSD-25 with an acceptable degree of success on cancellation of "squares," "vowels," "a-1," and "words." Prediction of ranks under 100 micrograms can be done from the zero dose ranks on cancellation of "squares," "a's," and "odd numbers preceded by odd numbers." One can predict ranks under 100 micrograms from those under 50 micrograms for cancellation of "squares," "a's," and "vowels."

#### REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psych.*, 1955, **39**, 3-60.
2. BUSCH, A. K., & JOHNSON, W. C. LSD-25 as an aid in psychotherapy. *Dis. Nerv. Syst.*, 1950, **11**, 2-4.
3. CONDRAU, G. Klinische erfahrungen an geisteskranken mit Lysergisaure-diethylamid. *Acta Psychiat. et Neurolog.*, 1949, **24**, 9-32.
4. DE SHON, H. J., RINKEL, M., & SOLOMON, H. Mental changes experimentally produced by LSD. *Psychiat. Quart.*, 1952, **26**, 35-53.
5. FORRER, G. R., & GOLDNER, R. D. Experimental physiological studies with lysergic acid diethylamide (LSD-25). *Arch. Neurol. & Psychiat.*, 1951, **65**, 581-588.
6. GARRETT, H. E. *Statistics in Psychology and Education*. New York: Longmans, Green, 1953. Pp. 451.
7. GASTAUT, H., FERRER, S., CASTELLS, C., LESEVRE, N., & LUSCHNAT, K. Action de la diethylamide de l'acide d-lysergique (LSD-25) sur les fonctions psychiques et l'electroencephalogramme. *Confina Neurol.*, 1953, **13**, 102-120.
8. HEBB, D. O. *The Organization of Behavior*. New York: Wiley, 1949. Pp. 335.

9. RINKEL, M., DE SHON, H. J., HYDE, R. W., & SOLOMON, H. Experimental schizophrenia-like symptoms. *Am. J. Psychiat.*, 1952, **108**, 572-578.
10. SAVAGE, C. Lysergic acid diethylamide (LSD-25); a clinical-psychological study. *Am. J. Psychiat.*, 1952, **108**, 896-900.
11. STERZINGER, O. Zerprufung und Untersuchung der abstrakten Aufmerksamkeit. *Psyche*, 1924, **23**, 121-161.
12. STOELTING COMPANY. Examiner Manual: Cancellation and Attention Tests. Chicago, Illinois. Pp. 3.
13. STÖLL, W. A. D-lysergic acid diethylamide (LSD-25), a phantasticum of the ergot group. *Schweiz. Archiv. für Neurolog. und Psychiat.*, 1947, **60**, 1-45.
14. THORNDIKE, E. L., & WOODWORTH, R. S. The influence of improvement in one mental function upon the efficiency of other functions. *Psychol. Rev.*, 1901, **8**, 247-261.
15. WHIPPLE, G. M. Manual of Mental and Physical Tests. Baltimore: Warwick & York, 1914. (Pp. 305-329.)
16. WILCOXON, F. Some Rapid Approximate Statistical Procedures. New York: American Cyanamid Co., 1949. Pp. 16.

*The Mount Sinai Hospital*  
*Fifth Avenue & 100th Street*  
*New York City 29*





## SEX STEREOTYPES AND ACCEPTANCE OF SEX ROLE\*

*Department of Psychology, Los Angeles State College*

---

SOLOMON DIAMOND

---

### A. PROBLEM

In our culture, as in every culture, certain ways of behaving are regarded as typically feminine, other ways as typically masculine. Opinions differ as to the extent to which such attitudes are based on sound observation, and also as to the nature of the influences which give rise to such uniformities as do exist. With these problems we shall not be concerned. We take as point of departure the fact that there are such sex stereotypes. We have tried to define experimentally certain aspects of these stereotypes, and to explore the possibility that the modification of these might be a projective expression of the individual personality. In particular, it seems reasonable to expect that "masculinity" and "femininity" will be somewhat differently perceived by women who accept their sex rôle and by women who reject it.

### B. METHOD

Seventy-nine students in a class in mental hygiene, and 20 in a class in the psychology of personality (early enough in the semester to be regarded as essentially naïve subjects) took what was ostensibly a "test of social perception." One part of this test consisted of an assignment to write a story about three characters, "a House, a Tree, and a Person." This material has been reported elsewhere (1), and is mentioned here only because it provides the basis for classifying our women subjects. Those who used feminine or juvenile persons in these stories are classified as accepting their sex rôle, while those who used adult male person or persons of indeterminate sex are classified as rejecting their sex rôle.

As another part of this "test of social perception," the subjects were asked to read a series of 60 statements, and to "indicate whether it is more likely that the speaker is masculine or feminine." A "?" response was also permitted, and was counted as a half choice for each sex. It was necessary to discard the records of two women who used this response almost exclusively, doubtless in a militant defense of sex equality.

---

\*Received in the Editorial Office on December 28, 1954, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

All probabilities stated throughout this paper are based on  $\chi^2$  values computed from four-fold tables. Since the cell entries frequently include half-units, because of the method of handling "?" responses, the correction for continuity was judged not applicable.

### C. RESULTS

#### 1. *The Stereotypes*

Men average 26.4 *M* and 26.0 *F* responses, while women average 27.8 *M* and 27.9 *F*. These differences are not significant.

With 60 items, there is a chance expectation that three items will yield  $P < .05$ , when testing the null hypothesis that the responses of men and women will be similar. This expectation is met exactly. We conclude that overall sex differences play a very minor rôle, and that we may therefore pool the responses of all subjects to arrive at composite judgments representing the stereotypes.

In view of the fact that the total numbers of *M* and of *F* responses are so nearly identical, we have tested each item against the null hypothesis that *M* and *F* responses would be equal. By this test, 19 items are indicated as "masculine" and 15 as "feminine" at the .001 level. One more is masculine, and three feminine, at the .01 level, and two more each at the .05 level. There remain only 18 items (including one on which men and women had significantly opposite judgments) with  $P > .05$ . It is clear that a consensus exists with respect to most items. Taken together, these constitute stereotypes.

We may indicate the general nature of these stereotypes, illustrating with a few of the statements on which general agreement prevails. Judged feminine are all indications of enthusiastic enjoyment (*Isn't that the most exciting thing that ever happened?*); of a dependent attitude (*What would I ever have done without you?*); of frustration resulting from enforced dependency (*I was never allowed to make up my own mind*); of a moody and withdrawn attitude (*I should be the happiest person in the world, yet somehow, I'm sad*); and most utterances of complaint (*Raining again! when will we ever have a little sunshine?*). Miscellaneous feminine items also include: *My father was the kindest man I ever knew. — Love is the only thing that gives meaning to life.*

Judged masculine are all statements that are boastful, or oriented toward goal attainment (*Just watch the way I do this!*); complaints about things that interfere with achievement (*I always get the toughest breaks*); rejection of advice (*Just leave me alone; I'd rather find out for myself*); and

offering it (*You'll know better when you've had as much experience as I*). Miscellaneous masculine utterances (which contrast clearly with the miscellaneous feminine items given above) include: *My mother was a wonderful woman. — The most important thing in life is a sense of humor.*

Items which were not predominantly judged as either masculine or feminine include those that are most strongly self-depreciatory (*What's wrong with me? I don't seem to be able to do anything worthwhile*); and some indicating very strong resentment (*My father was a tyrant in the family*).

## 2. Relation to Acceptance of Sex Rôle

Group  $W_m$  consists of 23 women who gave evidence of rejecting their sex rôle, as indicated in the first paragraph under "Method." Group  $W_f$  consists of 21 women who, by the same criterion, do not reject their sex rôle. (If the proportion of rejections seems high, it should be remembered that these are college students, most of whom are completing their educations in the face of material difficulties.) Group  $M$  consists of 51 male subjects. These groups will be compared in two ways: (a) by a correlation technique, and (b) by an item analysis technique.

a. *Correlation of the stereotypes.* For each group, the 60 items were ranked in order of increasing masculinity, and rank order correlations computed.  $M$  and  $W_m$  correlate .936;  $W_m$  and  $W_f$ , .919;  $M$  and  $W_f$ , .861. It can be shown (using Fisher's  $z$ -transformation) that the correlation between  $M$  and  $W_f$  is significantly less than between  $M$  and  $W_m$ . If the function  $1-\rho^2$  is used to represent "non-common variance," and comparing each group of women with the men, it appears that  $W_f$  is responding to twice as much "other influence" as  $W_m$ , in the ranking of items.

b. *Item analysis.* There are no items on which Groups  $M$  and  $W_m$  differ significantly.

Groups  $W_m$  and  $W_f$  differ significantly in their responses to two items. *What would you do in my place?* Group  $W_m$  gives about half  $M$  responses to this item, which is judged distinctly feminine by group  $W_f$ . This is perhaps only a matter of expressing the preferred identification, where the content has no obvious implication with respect to the  $M$ - $F$  variable. *I never have fun at parties.* Group  $W_m$  judges this to be feminine, while  $W_f$  judges it to be masculine. In other words, most members of  $W_m$  consider parties to be occasions when men have fun, while most members of  $W_f$  consider them to be occasions when women have fun.

Groups  $M$  and  $W_f$  differ significantly on six items. *What would you do in my place?* — *Please let me know if there is anything I can do to help.*



The greater likelihood of *W*<sub>1</sub> to call these feminine utterances may only represent readier identification. *I've never had trouble making friends wherever I go.* One feels that the acceptance of this as a feminine statement is linked with the acceptance of the feminine rôle. *This is the finest opportunity I could have asked for.* *W*<sub>1</sub> judges this even more strongly masculine than other subjects do, apparently expressing a disclaimer of career aspirations. *You'll know better when you've had as much experience as I.* — *I've always had to work hard for anything I got.* The greater readiness of *W*<sub>1</sub> to regard these as feminine statements (although even in this group the majority regard them as masculine) may indicate a disposition to view the mother rather than the father as occupying the rôle of mentor.

#### D. SUMMARY

Composite judgments by men and women college students, given in the course of what was ostensibly a test of social perception, show that almost any platitudinous remark is interpreted in terms of pervasive sex stereotypes. When women subjects are classified, on the basis of performance on another projective test, as either accepting or rejecting their sex rôle, it is found that the "rejecting" group holds sex stereotypes which do not differ significantly from those of men at any point, while the "accepting" group does show significant differences. These results support the basic hypothesis of this study, that modifications of popular stereotypes may be utilized as projective indices of individual personality characteristics.

#### REFERENCE

1. DIAMOND, S. The house and tree in verbal fantasy: I. Age and sex differences in themes and content. *J. Proj. Tech.*, 1954, 18, 316-325.

Department of Psychology  
Los Angeles State College  
855 North Vermont Avenue  
Los Angeles 29, California

## CHEMICAL INFLUENCES ON BEHAVIOR: I. THE EFFECTS OF A SMALL DOSE OF HYOSCINE ON PERFORMANCE\*

*Department of Psychology, University of Rochester*

S. J. ALEXANDER,\*\* M. COTZIN,† AND G. R. WENDT<sup>1</sup>

### A. INTRODUCTION AND SUMMARY

This study is concerned with the effects of small doses of hyoscine hydrobromide upon several kinds of performance. Doses of 0.6 mgm by mouth were given to young male adults for comparison with the effects of a placebo administered under the same conditions. An effort was made to control and counterbalance all significant variables affecting performance and affecting drug action. Eleven performance tests were selected to cover a variety of activities. Seven of the tests showed negligible differences between drug and placebo days (arithmetic, addition, same-opposites, disarranged sentences, analogies, complex coördinator); one (following directions) showed 5 per cent poorer performance on hyoscine days; three showed better performance on hyoscine days (rifle target fire 7 per cent, judgment 3 per cent, code substitution 3 per cent). Only the differences in following directions and in code substitution were statistically reliable at better than the 5 per cent standard. These results appear to be confirmation of the tentative conclusion (based on previous studies) that a single small dose of hyoscine by mouth does not damage performance of active men enough to contra-indicate its use.

### B. PURPOSE OF STUDY

Hyoscine in large doses is an obvious depressant of higher mental functions. Previous studies (1), motivated by an interest in the effects of small doses such as are used for prevention of motion sickness (0.6 mgm), and directed at finding a more or less dramatic effect of the drug on performance, had shown that such was not the case.

\*Received in the Editorial Office on January 9, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

\*\*Portland, Connecticut.

†Deceased.

<sup>1</sup>The work described in this paper was in part supported by a contract, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Wesleyan University, Middletown, Conn. Preparation of this report has been done in connection with a contract between the Office of Naval Research and the University of Rochester, contract N6ori-126, Task Order I. This is the first of a series of studies of effects of drugs done under OSRD, ONR, USPHS.

In none of the studies reported previous to the present one were the controls designed to be adequate to reveal small differences in performance caused by the drug. In view of the extensive use of hyoscine during war-time assault landing operations, it was decided to attempt a more closely controlled study of the effects of small doses upon performance.

### C. EXPERIMENTAL METHOD

#### 1. *General Plan of Experiment*

Each of 16 subjects took tests on each of six consecutive days,<sup>2</sup> starting on each day one hour after taking a capsule containing what was represented as a "medicine." On the first two days all subjects received milk sugar; on the last four days each subject had two days with milk sugar, two with 0.6 mgm of hyoscine hydrobromide. Only the data of the last four days were used for determination of drug effects. The types of test given and their sequence were identical on each day. Each testing period lasted 1½ hours. The procedure was designed to control all variables that were subject to control and to counterbalance the schedule of the experiment so as to equate the effects of those remaining variables (such as practice effect on tests and body weight of subjects) which could not be controlled. It was felt that this object was achieved except in the case of certain recognition-memory tests which were handled differently from the others.

A strategic problem of the experiment was to administer the maximum number of tests with the available personnel<sup>2</sup> and equipment of the project. The problem was solved by dividing the experimental group into two halves, each completely internally counterbalanced, using eight subjects from August 2-7, 1943; and the other eight from August 9-14, 1943. The groups of eight were subdivided (counterbalancing for body weight) into two groups arriving on a staggered time schedule, four of them at 3:00 P.M., four at 3:30. Within each of these groups, two were on one drug schedule, two on another. By using six laboratory rooms, with a person in charge of each, a shift system was devised whereby each subject moved to his next test on a time schedule set up for the whole (see Table 1 for details).

The tests were of two kinds: those administered individually by an assistant, and those administered to each subgroup of four as a group. The individual tests (complex coördinator, rifle target shooting, addition, code substitution) were each allotted seven minutes, including time to shift from

<sup>2</sup>In addition to the authors, the following aided in the conduct of the experiment: Mr. F. S. Apicella, Mr. R. M. Heath, Mr. E. A. Ricciuti, Mrs. Evelyn Ward.

room to room. The group tests (taken from the Army Alpha intelligence test and a recognition-memory test) were allotted 30 minutes.

Each group of four men received the capsule at the same hour and after 50 minutes assembled to complete a questionnaire on subjective symptoms. Each man then went to one of the four individual tests. After completing the four individual tests the group reassembled for the group tests. At the end of 30 minutes they again went to the individual tests and completed a second cycle in the same sequence as the original. Thus the session lasted  $1\frac{1}{2}$  hours during which each man took the individual tests twice, the group tests once. The staggered system of arrivals brought in the second group to take the individual tests just as the first group was going to the group tests. As the second group went to the group tests, the first returned to the individual tests. Finally, as the second returned to the individual tests, the first was leaving the laboratory.

The details of the problem of counterbalancing will not be discussed here. Study of the schedule displayed in Table 1 will show how this was accomplished. Several points may be noted which are not apparent from the table.

1. Two of the subjects of each group of four received the drug on Days 3 and 6, the placebo on Days 4 and 5, while the other two were on the opposite schedule; i.e., drug on Days 4 and 5, placebo on 3 and 6. The purpose of this schedule was to counterbalance practice effects and day to day variation due to weather, etc. The two initial placebo days were designed to reduce the variability of performance due to uncontrolled factors operating early in practice. Each man thus had four practice sessions on the individual tests before the experimental days began.

2. The individual tests were given to different men at different times (after taking the medication) within the two 30-minute periods devoted to these tests. The sequence of the four tests was thus different for each of the four men of a group. The first of these periods began one hour after the dose, the second two hours after the dose. The group tests were given simultaneously to four men in the 30 minutes beginning  $1\frac{1}{2}$  hours after the dose.

## 2. *The Tests*

The selection of tests should depend upon the nature of the experimental question. In the present case the question was unfortunately vague and too inclusive; i.e., does hyoscine damage performances of consequence and too inclusive? Tests should have been selected such that an individual who performs them well will also be able to perform any military task well. It is obvious to any who have experimented with performance



TABLE 1  
SCHEDULE OF TESTING

[illegible]

\*These subjects served during the second week; the others during the first week.

tests that such a criterion was impossible of attainment. One can make no special claims for the particular battery of tests used.

a. *Rifle target test.* The subjects fired indoors, using a .22 caliber, bolt action Savage Sporter rifle (7 lbs.) with a shaded blade front sight and peep rear sight, firing long rifle cartridges. Rapid-fire-type targets were used at 30 feet. During each test the subject fired five consecutive shots at each of two targets. Scoring was by measurement of the actual distance of each shot from the bull's-eye. Firing was at will from the standing position, using standard U. S. Army form. The first three sessions (six targets of five shots each) were devoted to instruction.

b. *The complex coördinator* (also known as the "Mashburn apparatus" or the Mashburn serial reaction test) was that used for selection of AAF pilots. By operating a rudder bar and manipulating a "stick" comparable to airplane controls the subject controlled the position of a light on each of three banks of lights so that these matched the position of other lights set by the apparatus. As soon as the subject achieved a complete match, the apparatus automatically presented a new pattern to be matched. The score was the total number of settings completed during five minutes.

c. *Miles' modification of the Johnson-Paschal code substitution test.* This is a paper and pencil speed test in which a square of 100 letters is translated into a 100-space square according to a code which requires the substitution of another letter of the alphabet for each code letter. The test was scored by time and errors. W. R. Miles had prepared 24 forms of the test, of which 13 were used in this study (one for an initial practice period, then one for each of the two daily sessions). This test was selected because of its apparent sensitivity to factors producing deterioration. Johnson and Paschal (3) found it to be one of the few tests showing early deterioration under anoxia.

d. *Addition test.* Numerous investigators of the effects of alcohol on performance have used continuous addition tests, finding this activity to be adversely affected by small doses of alcohol. For lack of an appropriate published form of the test, one was devised. The subject was required to add aloud certain 2-place numbers to columns of other 2-place numbers (the eighty 2-place numbers which do not end in zero) in haphazard orders. The problems appeared on a typed form arranged in groups of five. One of the four two-place numbers, (58, 65, 74, and 83), was added to five consecutive numbers in the group of 80, then another was added to the next five, another to the next five, and so on. The test was administered with a 3-minute time limit.

The reliability of these individual tests is indicated by the consistency with which individuals maintained their same relative positions at the different times they took the tests. The best computation to show this is the correlation between the average performances on Days 3 and 5 and the average performances on Days 4 and 6. Each average includes one placebo day and one drug day. The correlations are shown in Table 2. The test reliabilities shown are good; the rifle target test being lowest of the group.

TABLE 2  
RELIABILITY OF INDIVIDUAL TESTS

Test	Rank diff. correl. Rho	Pearsonian correl. <i>r</i>
Rifle target test	.94	.89
Complex coördinator	.97	.99
Code substitution*	.95*	.96*
Addition	.98	.99

\*Results of the code substitution test for the first week include some estimated scores. In the last session of the last day of Week I the printed sheets of Form 13 of the test contained 9 letters per line instead of 10, due to a printing error. This confused all subjects with resultant poor scores. The estimated score for Session 2, Day 6, was determined by adding to each subject's score for Session 1, Day 6, the average change for all 8 subjects between Session 1, Day 5, and Session 2, Day 5.

*e-j. Army Alpha group tests.* Tests of more complex mental performances than those involved in addition and code substitution were desired. Standardized tests with six forms of equal difficulty, one form for each day were needed. The only available instruments at the time were the Army Alpha-Wells Revisions, Forms 5, 6, 7, 8 and 9, and the Army Alpha Bregman Revision, Form B (4, 2)<sup>3</sup>. The last-named were used on Day 1, the others on the remaining five days. The standard instructions were used, but all time limits were cut to prevent the subjects from getting perfect scores (due to their high ability and to the effects of practice). The tests used and the time limits adopted after preliminary trial on other subjects are shown in Table 3. The table also shows test reliability as indicated by correlation of subjects' performances on Days 3 and 5 *vs.* 4 and 6. Scoring was in all cases by counting the total number of problems correctly solved in each test.

*k. Recognition-memory test.* It is commonly believed that large doses of hyoscine produce amnesia for the period of medication, so a "memory test" was included. There were two difficulties, neither of which was wholly satisfactorially

<sup>3</sup>Obtained from Psychological Corporation, N. Y. City.

TABLE 3  
ARMY ALPHA TESTS

	Time Limits		Reliability	
	Standard	Used by us	Rho	r
Arithmetical Problems	5 min.	2 min.	.84	.87
Practical judgments	1½ min.	1 min.	.84	.81
Synonyms-antonyms	1½ min.	1 min.	.92	.89
Disarranged sentences	2 min.	1 min.	.77	.83
Analogies	3 min.	1 min.	.77	.78
Following directions	4 min.	3½ min.	.76	.71

solved. First, there were no standardized memory tests available in the required number of equivalent forms. Second, it was impossible, in the same experiment, properly to control and counterbalance both the performance tests (described above) and the memory tests, where the retention of material from the day before is tested. A separate experiment would have been required for a properly counterbalanced memory test. For these reasons the present results are only approximations.

The test was a visual recognition-memory test. On each day (except the sixth), after the subjects had completed the Army Alpha group tests, they were shown 20 pictures to be learned. These were projected on a screen by means of a delineascope, one being presented every 10 seconds for a period of three seconds. Memory for them was tested on the next day. Just before taking the Army Alpha group tests, the 20 pictures of the preceding day were presented in a predetermined random order, mixed with 20 new pictures, selected for their similarity to the originals. The test series was again at the speed, one every 10 seconds. The subjects indicated by writing + or 0 whether they had seen each picture. (On Day 1 only a learning series was given; on Day 6, only a memory test.) Each test was for one type of picture: airplanes, flags of various nations, advertisements, military emblems, and wallpaper designs. Each picture was selected for its relative unfamiliarity. The difficulty of each series was roughly determined by preliminary experiments with other subjects. Table 4 shows the schedule of tests by days, the comparisons which could be made, and the number of subjects for which these comparisons were possible.

Reference to Table 4 shows that each of the possible sequences of placebo and hyoscine was used at least once. Exact comparison of recognition scores under each of these sequences would be possible if each of the five memory tests measured exactly the same activity. It is clear that this requirement is not fulfilled as indicated by lack of consistency of the subjects in the relative ranks on the five tests. Table 5 shows the rank order correlations



TABLE 4  
SCHEDULE OF RECOGNITION-MEMORY TESTS

Day	Schedule Week I	Schedule Week II	Performance comparison Conditions on presentation of picture	Conditions on test for memory	No. of subjects
1-2	Airplanes	Airplanes	Placebo	Placebo	16
2-3	Flags	Flags	Placebo Placebo	Placebo Hyoscine	8 8
3-4	Advertisements	Wallpapers	Placebo Hyoscine	Hyoscine Placebo	8 8
4-5	Emblems	Emblems	Placebo Hyoscine	Placebo Hyoscine	8 8
5-6	Wallpapers	Advertisements	Placebo Hyoscine	Hyoscine Placebo	8 8

(Rho) calculated from each subject's scores on each pair of tests. These correlations are based on scores according to the total number of correct recognitions. The errors (pictures shown but not recognized and pictures not shown but falsely recognized) were also scored for use in measurement of drug effect.

TABLE 5  
INTER-CORRELATIONS OF MEMORY TEST

	Wallpapers	Emblems	Advert.	Flags
Airplanes	.78	.74	.13	.31
Flags	.24	.36	.40	
Advertisements	.13	.59		
Emblems	.66			

### 3. Administration of Tests

The administration of all tests was under standardized conditions. Instructions were tested out in advance, typed, memorized, and then practiced by the assistants. The rooms used were free from distracting influences, and the whole procedure was accomplished without confusion or excitement.

### 4. Administration of Medication

Gelatin capsules were used, each capsule being half red, half white. The placebo was milk sugar; the drug was 0.6 mgm hyoscine hydrobromide. The capsule was taken with a half-glass of water in the presence of an assistant. No assistant knew the nature of the drug nor the schedule of medication, and all conversation about it or speculation concerning it was forbidden.

### 5. The Subjects; Control of Motivation

Sixteen civilian college students, average age, 18 years 2 months, average weight, 141 pounds, were used as subjects. They were carefully selected for above average grades, health, and emotional stability.

Good motivation and adherence to rules were achieved in the following ways. (a) In a one-hour group discussion the plan of the experiment, its importance, and its rules were explained. (b) Each subject signed a pledge to abstain from alcohol and excessive caffeine-containing beverages, to observe regular and specified schedules of sleeping and eating, to regulate smoking, to appear promptly, to exert himself to the utmost, to refrain completely from discussing the experiment with anyone (particularly with other subjects), and to refrain from studying or rehearsing the tests outside the Laboratory. Each day on arrival each subject made a signed statement con-

cerning these rules. (c) Each subject was given \$10 for the week on satisfactory completion of the study. (d) At pre-planned times the subjects were reminded of the desirability of full coöperation, playing on their patriotism, sense of duty, and interest in the task itself.

#### 6. *Control and Measurement of Suggestion Factors*

Much drug work is rendered doubtful in its interpretation because the subjects dramatize their rôles. Such effects are difficult to control satisfactorily because not all subjects react in the same way. It is therefore well to minimize suggestion factors and to induce a matter-of-fact attitude toward the medication. In this study the effort was to play down the medication but to play up the tests as the center of interest, while still making provisions for measurement and counterbalancing of suggestion effects. In the preliminary indoctrination the subjects were led to expect slight or undetectable effects. The dose was administered in a matter-of-fact manner. Then, the first two days (on which only milk sugar was used) served to convince the subjects that the "medicine" was not potent. By the third day, when hyoscine was used on some, they tended to report a symptom only if it were unequivocal.

The presence or absence of subjective effects was determined daily 50 minutes after medication.<sup>4</sup> The subjects assembled 10 minutes before the testing to fill out a questionnaire. The first 11 questions concerned details of eating, sleeping, etc., since the preceding day. The last 27 questions gave a complete opportunity to list symptoms. (Sample question: "Are your mouth and throat (a) drier than usual —, (b) normal —, (c) less dry than usual —?") This was done under the supervision of a person known to the subjects to be a physician.

The questionnaire showed that these subjects were mostly unaware of any symptoms 50 minutes after taking hyoscine, and that those symptoms which they did report were probably not the result of the drug. Reports of symptoms became progressively less during the six days. Each questionnaire listed 24 questions about particular symptoms. There were thus obtained 384 answers for each day (24 questions  $\times$  16 subjects). Each could be answered in three ways: no deviation from normal, deviation from normal in the direction of hyoscine effect, and deviation from normal in a direction not expected from hyoscine. One can call the last two "plus" and "minus" responses respectively. On the first two days (placebo) there were totals

<sup>4</sup>According to later studies the effects of hyoscine do not usually begin so soon as this (5).

of 33 plus and 24 minus responses out of a total possible 768. On the later placebo days (on Days 3, 4, 5, and 6) there were totals of 13 plus and 7 minus responses out of 768. On the actual drug days (on Days 3, 4, 5, and 6) there were only 9 plus and 7 minus responses (out of 768). The day-to-day totals of abnormal responses<sup>o</sup> (all plus and minus out of a possible 384) were 38, 19, 11, 7, 9, and 9. One question asked for an over-all judgment by the subject (To what degree do you think you now notice effects of the medication, (a) none —, (b) slight —, (c) moderate —, (d) considerable —, (e) extreme —?). This was actually checked only in the first two alternatives, "none" and "slight." The frequency (out of a possible 16 subjects) of checks of "slight" for each of the six days was 4, 3, 2, 2, 2, and 1. On the last four days three of the seven men reporting "slight" effects had received hyoscine, while four had received a placebo.

The conclusions seem justified that these subjects had few subjective effects due to suggestion and were unable to detect the drug 50 minutes after taking it. Reasons for this almost total absence of subjective effects are probably, (a) the smallness of the dose, (b) the counter-suggestion effects of the initial placebo days and of the experimental procedure where the subjects were kept continuously active and alert, (c) the difficulty of detecting effects only 50 minutes after taking the dose.

#### D. RESULTS

##### 1. *Army Alpha Tests*

Table 6 shows that the differences between drug and placebo days were negligible (less than 1 per cent) on tests of arithmetic, same-opposites, disarranged sentences, and analogies. The judgment test scores were about 3 per cent better on drug days, while scores for following directions were about 5 per cent better on placebo days.

TABLE 6  
EFFECTS OF HYOSCINE ON ARMY ALPHA TESTS

Test	Percentage differences between drug and placebo days based on total number of correct answers		
	+ favors drug	— favors placebo	
	First eight subjects	Second eight subjects	All sixteen subjects
Arithmetic	—1.1%	0.0%	—0.5%
Judgment	+6.1%	0.0%	+2.8%
Same-opposites	—0.2%	+0.8%	+0.3%
Disarranged sentences	—3.9%	+5.6%	+0.9%
Analogies	—3.7%	+3.4%	—0.2%
Directions	—4.6%	—5.0%	—4.8%



Estimation of the reliability of differences by the formula  $\sigma_{diff.} =$

$\sqrt{\sigma_{M_p}^2 + \sigma_{M_h}^2 - 2r\sigma_{M_p} \cdot \sigma_{M_h}}$ , followed by calculation of critical ratios and of probability values leads to the conclusion that these differences are not statistically reliable except for the directions tests whose  $P$  value was .02. Calculations using the error scores showed that scores taking account of wrong answers leave this conclusion unaffected.

## 2. Individual Tests

Table 7 shows results in terms of percentage differences between drug and placebo days. The difference between drug and placebo days was negli-

TABLE 7  
EFFECTS OF HYOSCINE ON INDIVIDUAL PERFORMANCE TESTS

Test	Percentage differences between drug and placebo days		
	+ favors drug	- favors placebo	
	First eight subjects	Second eight subjects	All sixteen subjects
Rifle target test	+7.7%	+5.8%	+6.7%
Complex coördinator	-2.8%	-0.4%	-1.6%
Code substitution	+5.6%	+1.1%	+3.3%
Addition			
Total attempted	+1.6%	-0.4%	+0.6%
Total correct	+0.9%	-0.5%	+0.3%

gible on the addition test, whether scored by total problems attempted or by total correctly solved; and that the complex coördinator showed only a 1.6 per cent difference. Code substitution was better on drug days by 3.3 per cent and the accuracy of shots was better on drug days by 6.7 per cent. Only one of the latter differences can be regarded as a reliable index of a drug effect, as indicated by  $p$  values as follows: code substitution,  $p = .04$ ; rifle target test,  $p = .11$ .

The same data were analyzed to find whether the drug had a greater effect early in the afternoon or late in the afternoon, by comparing performance in the session at the beginning of the second hour and the session at the beginning of the third hour. The second tests of the afternoon on the whole yielded a better score than the first tests on both drug and placebo days. The degree of improvement from the first test to the second test was, however, greater on the days when hyoscine had been given, for each of the four tests.

The rifle target test  
showed

3.5 per cent more improvement on hyoscine days.

The complex co-  
ordinator showed

0.4 per cent more improvement on hyoscine days.

The code substitution test showed	0.7 per cent more improvement on hyoscine days.
The addition test showed	1.4 per cent more improvement on hyoscine days.

The magnitude of these differences is obviously so small as to be incapable of support by tests for statistical significance, deriving plausibility only from its consistency. If the change be granted as true it could be attributed to a recovery from depressant aspects of the drug effect between the second and third hours, were it not that other studies indicate that the maximum action of hyoscine occurs two to three hours after ingestion (5).

### 3. Recognition-Memory Tests

The purpose of the recognition-memory tests was to study drug effects upon the ability to learn visual material and upon the ability to recognize it 24 hours later, when learning and recognition were under varying conditions of medication. For reasons stated above, this part of our experiment is not an adequate test either of "memory" or of the effects of hyoscine on it. One had to depend partly on comparisons based on giving different memory materials to different subjects. The low correlations between the tests (Table 5) indicate the impropriety of this. Inspection of all of the data shows no obvious drug effects. Full presentation of these data seems unjustified in view of their inadequacy. The display is restricted to two of the best-controlled comparisons: (a) learning under placebo, recognition under placebo compared with learning under hyoscine, recognition under hyoscine (*PP vs. HH*) and (b) learning under placebo, recognition under placebo compared with learning under placebo, recognition under hyoscine (*PP vs. PH*).

a. *PP vs. HH*. On Days 4-5 during both weeks the test materials were pictures of emblems (see Table 4). On these days half of the subjects were on *PP* and the other half on *HH*. The average number of emblems correctly recognized by the eight *S*'s under *PP* was 36.8 and by those under *HH* was 35.3, a difference of 1.5 pictures in favor of placebo. But the *HH* group was probably a somewhat inferior group of subjects as indicated by comparisons of total pictures correctly recognized during the experiment: *PP* = 34.7 pictures per subject, *HH* = 33.8 pictures per subject. It therefore seems proper to conclude that no effect of hyoscine has been shown.

b. *PP vs. PH*. On Days 2-3 the test pictures were flags. On these days half of the *S*'s were on the *PP* schedule, half on *PH*. The scores for pictures correctly recognized were: *PP* = 33.8, *PH* = 34.0.

These data, and such other less-controlled comparisons as it was possible to make, give inadequate ground for concluding that hyoscine affected visual recognition memory.

### E. DISCUSSION

According to the results of this study 0.6 mgm of hyoscine hydrobromide taken by mouth probably has negligibly small effects on performance. On the whole, more types of performance were improved by the drug than were damaged.

This conclusion should not be generalized to apply to the effects of hyoscine on all tasks. Since all of these tests required continuous activity and were done under high motivation, the results may not apply to tasks which are not well-motivated. Nor did the tests include tasks requiring initiative and planning. The conclusions should not be generalized to bigger doses. This drug is quite toxic in large doses. The fact that depressant effects are absent after small doses is not an indication that they would also be absent after larger doses. Nor is the absence of an effect from a single dose proof that it would also be absent after repeated doses.

### REFERENCES

1. BARD, P., & TYLER, D. B. Motion Sickness. *Physiol. Rev.*, 1949, **29**, 311-369.
2. BREGMAN, E. O. Revision of Army Alpha Examination, Form B, Manual of directions. New York: Psychological Corporation, 1935.
3. JOHNSON, H. M., & PASCHAL, F. C. Psychological effects of deprivation of oxygen—deterioration of performance as indicated by a new substitution test. *Psychobiology*, 1920, **2**, 193-237.
4. WELLS, F. L. Army Alpha revised. *Person. J.*, 1932, **10**, 411-417.
5. WENDT, G. R., & YORBURG, L. Studies of motion sickness remedies: II. The speed and duration of action of hyoscine hydrobromide. (unpublished.)

*Department of Psychology*  
*University of Rochester*  
*Rochester 3, New York*

# STUDIES OF MOTION SICKNESS: X. EXPERIMENTAL PROOF THAT AVIATION CADETS TELL THE TRUTH ON MOTION SICKNESS HISTORY QUESTIONNAIRES\*<sup>1</sup>

*Department of Psychology, University of Rochester*

S. J. ALEXANDER, M. COTZIN, C. J. HILL, JR., E. A. RICCIUTI, AND  
G. R. WENDT\*\*

## A. INTRODUCTION AND SUMMARY

This report gives a research answer to the problem of whether aviation cadets tell the truth on questionnaires concerning their history of motion sickness. A common objection to their use has been that answers cannot be trusted because it appears in the interest of the cadet to falsify. Questionnaires were given, under conditions which could be duplicated in the armed services, to 349 naval aviation cadets shortly after arrival at flight preparatory school. One to 10 weeks later they were interviewed under optimal conditions of cooperation to determine their face-to-face answers to the same items as had appeared on the questionnaire. An analysis was made of the 4,883 recorded pairs of answers to find the kinds of changes. The results show that 80 per cent of the answers were identical; 9 per

\*Received in the Editorial Office on January 9, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

\*\*Alexander: Portland, Connecticut; Cotzin: deceased; Hill: Lawrence College, Appleton, Wisconsin; Ricciuti: 18 Forest Ridge, Waterbury, Connecticut.

<sup>1</sup>The work described in this paper was done under contracts, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Wesleyan University, Middletown, Connecticut, and the University of Rochester, Rochester, New York. The preparation and distribution of this report were supported by a contract between the Office of Naval Research and the University of Rochester (N6ori-126, Task Order 1). The questionnaire used (MSI-Form XP5, NRC-CAA-CSTAP-Item II-E-1942) had been developed by G. R. Wendt at Wesleyan University with the aid of funds furnished by the Division of Research, Civil Aeronautics Administration through the National Research Council's Committee on Selection and Training of Aircraft Pilots and was used in this investigation with their permission. This is one of a number of studies of motion sickness questionnaires performed under the auspices of the Committee on Selection and Training of Aircraft Pilots, the Committee on Medical Research and the Office of Naval Research under the general direction of G. R. Wendt. The subjects of the present investigation were naval aviation cadets and the data were collected by special permission of the Secretary of the Navy. The investigators wish to express their appreciation to the universities for making facilities available, to the sponsoring agencies for agreeing jointly to support the studies, to the naval officers whose cooperation and active assistance made the study possible and to the cadets who furnished the data.



cent changed from a questionnaire-answer "not enough experience" to an interview-answer "never sick," and 1 per cent *vice versa*; 1 per cent changed from a questionnaire-answer indicating sickness to an interview-answer indicating lack of experience, or *vice versa*; 5 per cent changed the interview-answer toward a greater degree of sickness, and 4 per cent changed the interview-answer toward a lesser degree of sickness. Changes of a magnitude large enough to be possibly indicative of falsehood occurred in 0.6 per cent of the answers. But of these 29 changes, some or all are better explained by errors of entry on questionnaire or interview, differences in judgment as to whether a particular case of sickness was due to motion or to other causes, and instances of sickness which occurred between the time of questionnaire and interview. It is concluded that this study gives no support whatever to the allegation that aviation cadets may be expected to falsify questionnaire-answers.

## B. PROCEDURES

### 1. General

The data described in this paper were collected as part of a larger program of studies of airsickness, laboratory motion sickness, and of personal history questionnaires for measurement of susceptibility to motion sickness. This is one of a series of reports on the reliability and validity of the questionnaire (2, 3, 5, 6, 7). Report No. 1 in this series (1) described the general techniques.

### 2. Subjects

The subjects were 349 naval aviation cadets resident at the Wesleyan University Naval Flight Preparatory School in 1944. The interview data were collected in the Laboratory at the time the subjects were going through other experimental procedures. Although service as subject was voluntary, we obtained a complete sampling of subjects because the selection was made by us and none declined to serve.

### 3. The Questionnaire

The personal history questionnaire is included in a preceding report (3). It contains 42 questions of which Nos. 10 to 23 concern motion sickness in boats, autos, trains, cars, buses, elevators, hammocks, lawn swings, merry-go-rounds, ferris wheels, roller coasters, the whip, airplane gliders or loop-o-planes, and airplanes. In the case of the first five devices the subjects were required to give one of five answers to each:

1. Have not had enough experience to know whether I would be affected. 2. Have never been affected by rides. 3. Have been somewhat subject to nausea, but not to the point of vomiting. 4. Used to be subject to sickness (a few years ago or as a child) but am not now subject to it. 5. Am now subject to sickness. Would probably vomit if the ride were long and rough.

In the case of the other devices one of four answers was required for each question:

1. I have not had enough recent experience to know whether I would be affected. 2. I have had experience, and have been unaffected by rides. 3. I have been slightly nauseated. 4. I have been strongly nauseated (to the point of vomiting or almost to it).

These 14 questions appeared, it should be noted, in the middle of the questionnaire, surrounded by personal history and medical history questions. Answers were entered on IBM machine-scored answer sheets.

The questionnaires were administered a few days after the Cadets entered the Navy. They were assembled in groups on order of the Commanding Officer. There they were told that they were to take a "special questionnaire for the medical records" and it was further explained that, "The information is asked so that your medical officers can keep you in good condition." It was administered by civilian personnel. It should also be noted that these Cadets took this questionnaire after having undergone a long preceding series of tests and questionnaires during selection, enlistment, and indoctrination. They were cooperative but unexcited and relatively inquisitive.

#### 4. *The Interview*

One to 10 weeks later selected cadets were invited to serve as subjects on the "Wave Machine" in the "Aeromedical Research Laboratory." There they took performance tests, were exposed to motion on the vertical accelerator, again took performance tests, and were then interviewed. The whole procedure took about 50 minutes.

The interviews were conducted by four experimenters (Alexander, Cotzin, Hill, and Ricciuti) according to a standard procedure. Its objective was that of finding the circumstances leading the subject to report himself as susceptible or otherwise. Answers were entered on a printed form, our Motion Sickness Inventory of December 1941 (4). The interviewer was the judge of where an answer should be entered. The atmosphere of the interview was warm, friendly, and cooperative, being guided toward a physician-and-patient relationship. Since the continued cooperation

tion of the cadets was essential to the conduct of our work, we had put a great deal of effort into the development of a procedure that would produce friendly relations.

### 5. *Assumptions of the Experiment*

The validity of this experiment as a test of the proposition that aviation cadets will lie about their history of sickness depends upon the validity of our assumptions. The following were our assumptions: (a) That, as volunteers and as newly enlisted men, the cadets wanted to stay in flight training. (b) That falsification might occur at the time of questionnaire because the cadets received a minimum of information about its purpose, received a somewhat ambiguous explanation which suspicious men could distrust, and were strangers to the personnel who administered it, hence without personal loyalties. (c) That the interview-answers were made without the aid of significant recall of the questionnaire-replies, since the original questionnaire contained many items and was but one of many they had taken, and one to 10 weeks had intervened. (d) That the interview-answers were in all or nearly all cases made with full intent to tell the truth.

### 6. *Treatment of Data*

The total possible pairs of answers was 4,886 ( $349 \times 14$ ). Due to omissions, only 4,883 were recorded. Analysis consisted of counting the frequencies of instances where the questionnaire- and interview-answers were identical for each of the 61 items on the questionnaires, and the frequencies of each of the 208 possible kinds of changes from questionnaire-answer to interview-answer. For ease of presentation these data are grouped into somewhat larger classifications. (No analysis of individual subjects was made. Another report in this series (5) deals with such analyses.)

## C. RESULTS

The results of this analysis are shown in Table 1. It will be noticed that 80.2 per cent of all answers at the time of interview were identical with the answers given on the questionnaire. These cases of no change were divided as follows: 89.9 per cent were answers of "no sickness," 5.7 per cent were answers of "no experience," 3.5 per cent were answers of "slight sickness or childhood sickness," and 0.9 per cent were answers of "now susceptible."

Changes involving judgments of "not enough experience" occurred in 11 per cent of the pairs of answers. Of these, 9.5 per cent answered "no experience" on the questionnaire but were changed by the interview (to

"no sickness" or others) and 1.7 per cent went the other way. Only 0.7 per cent of pairs were in a direction which might indicate evasion at the time of the questionnaire; i.e., questionnaire-answer of "no experience" but interview-rating of sickness.

TABLE 1  
CHANGES IN ANSWERS FROM QUESTIONNAIRE TO INTERVIEW

Comparison of interview-answer and questionnaire-answer	Number of pairs of answers	Per cent of pairs of answers
No change of statement	3,916	80.2%
Changed from "No experience" to "No sickness"	429	8.8
Changed from "No experience" to Sickness	36	0.7
Changed from "No sickness" to "No experience"	65	1.3
Changed from "Sickness" to "No Experience"	18	0.4
Total changes from or to "No experience"	548	11.2%
Changed one step toward more sickness	140	2.9%
Changed two steps toward more sickness	82	1.7
Changed three steps toward more sickness	8	0.2
Total changes toward more sickness	230	4.7%
Changed one step toward less sickness	165	3.4%
Changed two steps toward less sickness	24	0.5
Changed three steps toward less sickness	0	0.0
Total changes toward less sickness	189	3.9%
Grand totals	4,883	100%

Changes from a lower level of sickness to a higher level occurred in 4.7 per cent of pairs of answers, while changes from a higher to a lower level occurred in 3.9 per cent of pairs. The difference between these two values is 0.8 per cent, which represents the excess of changes in the direction expected if falsification took place at the time of the questionnaire.

On the questionnaire, 3,730 answers were recorded as "no sickness." This is 76.3 per cent of all answers. Their interview-answers were divided as follows: 94.5 per cent again gave the answer "no sickness," 3.8 per cent said "slightly sick or now susceptible," and 1.7 per cent changed to the answer "no experience."

Large changes from questionnaire to interview, difficult to attribute to differences in attitude at the time of questionnaire and interview, are 3-step changes for the first five questions and 2-step changes for the last nine. On counting these 2- and 3-step changes, we find 29 (0.6 per cent) changes toward greater sickness in the interview report and 3 (0.1 per cent) toward less sickness. The difference between these two values is 26, which is 0.5 per cent of the total pairs of answers.



## D. DISCUSSION

The total frequencies of changes are thus seen to have been relatively small and the frequencies of large changes almost vanishingly small. It is, at any rate, too small to be of significance in the practical application of questionnaires to large groups of subjects.

The causes of the changes were found to include the following: (a) A difference in interpretation of what constitutes "enough experience to know whether I would be affected." (b) A difference in opinion as to whether they are "now somewhat subject to sickness," especially as affected by the fact that many of them had just come off the vertical accelerator and had been sick. (c) Errors of entry on the machine-scored answer sheet, made at the time of questionnaire. (d) Errors of entry by the interviewer. (e) Differences of opinion between cadet and interviewer as to whether a given instance of sickness was due to motion or to other causes such as alcohol, infectious disease, gastric upsets, etc. (f) Instances where the cadets took boat-rides, bus-rides, airplane rides, etc., in the one to 10 weeks between questionnaire and interview.

When all of the above causes are taken into account, falsification on the questionnaire as a cause of changes from questionnaire to interview must be regarded as an exceedingly rare or perhaps absent occurrence.

The achievement of a testing-situation for administration of questionnaires such that men will tell the truth appears not to be particularly difficult. We conjecture that the main requirement is that they be given enough questionnaires to be somewhat incurious about each new one. We have no doubt, however, that questionnaires could be administered in other contexts and under other conditions with a result quite different from this one.

## REFERENCES

1. ALEXANDER, S. J., COTZIN, M., HILL, C. J., RICCIUTI, E. A., & WENDT, G. R. Wesleyan University studies of motion sickness: I. The effects of variation of time intervals between accelerations upon sickness rates. *J. of Psychol.*, 1945, 19, 49-62.
2. ———. Wesleyan University studies of motion sickness: VI. Prediction of sickness on a vertical accelerator by means of a motion sickness history questionnaire. *J. of Psychol.*, 1945, 20, 25-30.
3. COTZIN, M., HILL, C. J., & WENDT, G. R. Studies of motion sickness: IX. The prediction of airsickness and of laboratory motion sickness by means of a personal history questionnaire. (Author complete.)
4. LINDSLEY, D. B., & WENDT, G. R. An investigation into the relationship of the electroencephalogram to motion-sickness susceptibility. In *Studies in motion sickness*, Series A. Division of Research, Civil Aeronautics Administration, Report No. 40, December, 1944.
5. WENDT, G. R. Studies of motion sickness: XI. The test-retest reliability of a motion sickness history questionnaire. In *Committee on Aviation Medicine, Final Report*, National Research Council, 1945.

6. ———. Airsickness among one hundred eighty-nine airline stewardesses and its relationship to previous history of motion sickness. *In* Studies of motion sickness, Series B. Division of Research, Civil Aeronautics Administration. Report No. 60, April, 1946.
7. ———. Airsickness among seventy-one student pilots and fifteen instructors and its relationship to previous history of motion sickness. *In* Studies of motion sickness, Series B. Division of Research, Civil Aeronautics Administration. Report No. 60, April, 1946.

*Department of Psychology*  
*University of Rochester*  
*Rochester 3, New York*



## STUDIES OF MOTION SICKNESS: XIII. THE EFFECTS OF SICKNESS UPON RIFLE TARGET SHOOTING\*<sup>1</sup>

*Department of Psychology, University of Rochester*

S. J. ALEXANDER, M. COTZIN, J. B. KLEE, AND G. R. WENDT\*\*

### A. INTRODUCTION AND SUMMARY

Tests of rifle fire accuracy at 25 feet from the target have been administered to 117 subjects before and after exposure to motion on our vertical accelerator. The data were analyzed to determine whether those who became motion sick showed deficits in performance. The average deviation from the target center of shots by 26 sick subjects showed a mean deterioration of 0.1 mm in performance, whereas 91 who did not become sick showed a mean improvement of 1.6 mm. This difference is not statistically reliable, as shown by a critical ratio of only 0.5. Records of speed of firing indicated no tendency for sick subjects to slow down after wave-machine exposure. It was found that the sick subjects, when asked to predict their own post-exposure performance, tended to drop their estimates further below their pre-exposure performance than did the not-sick subjects. Other analyses showed that neither susceptibility to machine sickness nor past history of sickness bore any demonstrable relation to normal performance on the tests administered before exposure to motion.

### B. PROCEDURES

The present paper presents performance test results which were recorded collaterally to other observations, elsewhere reported (4). The general procedures used have been fully described (1, 2) and will not be repeated here. The subjects were 86 Naval Officer-Candidate Trainees and 31 male civilian college students. They were exposed to motion on a vertical accelerator with rifle target shooting tests before and after. For

\*Received in the Editorial Office on January 9, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

\*\*Alexander: Portland, Connecticut; Cotzin: deceased; Klee: Brandeis University, Waltham, Massachusetts.

<sup>1</sup>The work described in this paper was done under contracts, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Wesleyan University, Middletown, Conn., and the University of Rochester, Rochester, New York. The preparation and distribution of this report were supported by a contract between the Office of Naval Research and the University of Rochester (N6ori-126, Task Order 1).



the purposes of the present report the protocols of the subjects were divided into three groups: the not-sick, the slightly sick, and the vomiters. They were compared in terms of the changes which took place in their performance before and after exposure to motion.

Each subject fired five shots at each of eight targets: one target purely for instruction and demonstration, followed by one target for practice; then a three-minute rest followed by three test targets separated by one-minute rests; then exposure on the wave-machine followed by a three-minute rest; finally three more test targets separated by one-minute rests.

The rifle was a .22 caliber, 7-pound Savage Sporter with a fine-aperture rear peep-sight and a shaded front blade sight. It was fired according to standard U. S. Army form from the standing position (except that no strap was used). The target distance was 25 feet, being so limited by the available space. The subjects were told to complete the five shots at each target in one minute, and were reminded after each target whether they were slow or fast, but no effort was made to force them to conform. Times were recorded, unknown to the subjects.

Before firing at a fresh target, the subject was told his score on the preceding target and was required to estimate his score on the next target. Both scores were recorded. For this purpose the printed scoring instructions on the targets were used. The maximum score was 25. From these we calculated the discrepancy of the predicted scores from the scores of the preceding targets on which they were based.

For the purposes of assessing the effects of sickness, accuracy was measured by the actual distance of shots from the target center. An exception was that all shots striking more than 10 cm from center were counted as having struck at 10 cm. This procedure was adopted after making a frequency distribution of the distance of all 4,680 shots. They fell into a skewed distribution whose mode was at 15 mm from the target center, whose median was at 16 mm, whose mean was 21 mm with an *SD* of 11.5 mm, and whose rather flat tail between 10 cm and 28 cm contained only 2 per cent of the total shots.

The reliability of the rifle test was estimated from the pre-exposure test performance and post-exposure test performance of the 91 subjects who did not get sick. If we make the assumption that the wave-motion affected all of the not-sick alike, so far as the test is concerned, the correlation of their two sets of scores is a proper test of reliability. This correlation was  $r = +.71$ , which indicates only moderate reliability.

## C. RESULTS

The effects of motion sickness upon accuracy are displayed in Table 1, where the scores of the sick can be compared with those of the not-sick. It will readily be seen that the changes in performance from pre-exposure

TABLE 1  
EFFECTS OF SICKNESS UPON RIFLE ACCURACY  
(Mean distance from bull's-eye in mm)

Degrees of sickness	No. of cases	Pre-test Mean	Post-test Mean	Change* Mean	SD
Not sick	91	20.3	18.7	+1.6	6.1
Slightly sick	19	20.7	21.0	-0.3	6.8
Vomited	7	19.4	18.9	+0.5	6.4
SS + V	26	20.4	20.5	-0.1	6.7

\*(-) = deterioration in performance; (+) = improvement.

test to post-exposure test are small for all groups, whether sick or not sick. In no case is the amount of change reliably greater than zero. The critical ratios of the differences between pre- and post-tests are all 0.4 or less in each group. If we then go on to compare groups (i.e., sick with not-sick), the differences in the relative amounts of change in the sick as compared to the not-sick are also not reliably greater than zero. The largest critical ratio is 0.5. We may therefore conclude that we have failed to demonstrate any effect of wave-machine exposure *per se* or of motion sickness *per se* upon rifle target shooting accuracy.

Our data on accuracy are somewhat more fully displayed in Table 2, which gives the average performance on each of the eight consecutive tar-

TABLE 2  
CHANGES IN ACCURACY ON CONSECUTIVE TARGETS  
(Mean distance from bull's-eye in mm)

Degree of sickness	No. of cases	Practice targets		Pre-tests			Post-tests		
		1	2	1	2	3	1	2	3
0	91	24.4	22.5	20.9	20.3	19.7	19.0	18.8	18.3
Sl. Sick	19	23.7	23.7	21.9	20.2	20.1	22.1	20.8	20.3
Vomit	7	19.3	18.3	17.6	21.6	19.0	21.7	16.9	18.0

gets for each sickness group. This material is included for the use of those who may be interested in progressive changes in scores.

The data on speed of firing are displayed in Table 3. The main results shown by this table are a small increase in speed with early practice and a marked steadiness of time during performance. From this steadiness we con-

clude that these data cannot be used to account for the failure to find deterioration in accuracy as a result of sickness.

TABLE 3  
EFFECTS OF SICKNESS UPON SPEED OF FIRING  
(Seconds to fire five shots)

Degree of sickness	Practice targets		Pre-tests			Post-tests		
	1	2	1	2	3	1	2	3
0	59	52	49	49	49	46	46	46
Sl. sick	67	54	54	51	51	48	44	46
Vomit	57	46	47	42	46	48	42	48

Data on the subjects' predictions of their target scores are displayed in Table 4. For the purpose of these predictions the targets were scored according to the score system printed on them. A perfect score was 25; 3 was the lowest score any of the present subjects actually obtained. Positive numbers in Table 4 mean that the subjects predicted that their next target would be better than their preceding target; negative numbers mean that they predicted a loss. The data are shown for each sickness group. The detailed discussion of this table is outside the province of this paper. It is presented here merely to show that there was a tendency for the sick subjects to predict that their sickness would produce a deterioration of performance. (Compare pre-test Target 3 and post-test Target 1.) But this prediction of deterioration was made only on the first target after sickness, and was altered for the two subsequent targets.

TABLE 4  
DISCREPANCIES BETWEEN ESTIMATED FUTURE ACCURACY AND PAST PERFORMANCE

Degree of sickness	No. of cases	Pre-test targets			Post-test targets		
		1	2	3	1	2	3
0	91	+0.9	+0.8	+0.5	-0.6	0.0	+0.5
Sl. Sick	19	+1.2	+0.8	-0.1	-2.5	-0.1	-0.1
Vomit	7	+0.7	+1.0	+2.1	-1.4	+1.0	0.0

There is one final relationship that is worth pointing out: the relationship of past history of motion sickness as determined by questionnaire (2) to the rifle scores. This relationship helps to answer the question whether "susceptible" individuals differ in original skill or otherwise from "non-susceptibles." The present group of subjects contained 31 classified as "non-susceptible," 43 "slightly susceptibles," and 43 "susceptibles." Their average pre-exposure scores were respectively 20.1 mm, 19.6 mm, and 21.0 mm. Since

these do not differ by amounts of any consequence, and since they reveal no trend, it may be concluded that no relationship between susceptibility and skill is demonstrated. Reference to Table 1 shows that when machine sickness (i.e., sickness in this experiment) is compared to skill, the same lack of relationship is apparent. The pre-exposure scores were: not-sick, 20.3 mm; slightly sick, 20.7 mm; vomited, 19.4 mm.

#### D. DISCUSSION

The present results are consistent with those presented in a previous report (3) in showing unimportant effects of motion sickness upon laboratory performance tests. In that report we warned against the direct application of such a finding to the military situation. The considerations there presented hold equally well for the present study.

#### REFERENCES

1. ALEXANDER, S. J., COTZIN, M., HILL, C. J., RICCIUTI, E. A., & WENDT, G. R. Wesleyan University studies of motion sickness: I. The effects of variation of time intervals between accelerations upon sickness rates. *J. of Psychol.*, 1945, **19**, 49-62.
2. ———. Wesleyan University studies of motion sickness: VI. Prediction of sickness on a vertical accelerator by means of a motion sickness history questionnaire. *J. of Psychol.*, 1945, **20**, 25-30.
3. ———. Wesleyan University studies of motion sickness: VII. The effects of sickness upon performance. *J. of Psychol.*, 1945, **20**, 31-39.
4. ALEXANDER, S. J., COTZIN, M., KLEE, J. B., & WENDT, G. R. Studies of motion sickness: XVI. The effects upon sickness rates of waves of various frequencies but identical acceleration. *J. Exper. Psychol.*, 1947, **37**, 440-448.

Department of Psychology  
University of Rochester  
Rochester 3, New York





## STUDIES OF MOTION SICKNESS: XIV. SUBJECTIVE REPORTS OF THE APPARENT PATH OF MOTION ON A VERTICAL ACCELERATOR\*<sup>1</sup>

*Department of Psychology, University of Rochester*

M. COTZIN, C. J. HILL, JR., AND G. R. WENDT\*\*

### A. SUMMARY

Subjects exposed to wave motion on a vertical accelerator were asked to describe the apparent path of motion. Out of 267 subjects, 93 reported only one kind of motion, 122 reported two kinds, 44 three kinds, 6 four, and 2 five kinds. Most of the reported motions other than the true vertical were ellipses or modified ellipses.

### B. INTRODUCTION

The observations here reported were made collaterally to a series of experiments on the effects of wave-form upon motion sickness in young men exposed to the up and down motion of a vertical accelerator. The present problem arose from earlier experiments during which occasional subjects expressed puzzlement concerning the path of the motion. They reported that it felt to them as though they were moved in an elliptical path (or others) and as though the character of the motion changed during the session.

These perceptual variants from the true path seemed to us to merit further investigation in an exploratory way. We did not have any definite hypotheses in mind, but it seemed possible that the results might shed light on the physical phenomena occurring in the vestibular apparatus during vertical acceleration or on some unsuspected conditions of motion sickness.

### C. PROCEDURES

#### 1. General

The general procedures of these experiments have been described in Report I of this series. The specific procedures used on the present subjects

\*Received in the Editorial Office on January 9, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

\*\*Cotzin: deceased; Hill: Lawrence College, Appleton, Wisconsin.

<sup>1</sup>The work described in this paper was done under contracts, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Wesleyan University, Middletown, Connecticut, and the University of Rochester, Rochester, New York. The preparation and distribution of this report were supported by a contract between the Office of Naval Research and the University of Rochester (N6ori-126, Task Order I).

are described in Reports IV and XVI. For the purposes of the present paper it is sufficient to note that the subjects came to the Laboratory individually for a 50-minute session during which they took certain performance tests, were then exposed to the motion of the vertical accelerator for 20 minutes (they were taken off sooner if they vomited), and were then again given the performance tests.

## 2. *The Motions*

The vertical accelerator is like an elevator cab. Its motions were produced by reversing the direction of acceleration at the instant the cab reached its full upward or full downward velocity so that it was in continuous up and down motion. Each subject was seated, the ear-eye line horizontal, the head unsupported, and blindfolded. The waves used are described in detail elsewhere. They were those designated by the code-letters H', J', K, L, M, N, P, Q, and R. Waves K, L, M, and N were asymmetrical in that the acceleration in the top half of the wave was different from that in the bottom half. The H', J', P, Q, and R waves were symmetrical but varied from one another in acceleration, amplitude, and frequency. See Reports IV and XVI.

## 3. *Instructions and Protocols*

We wished to strike a compromise between undue restriction and undue freedom of subjective report. Instructions to be prepared to report on the apparent path of motion were given to each subject immediately before being put in the cab. The subject was briefly shown a series of diagrams, Figure 1, and told: "The wave machine travels straight up and down. Some people perceive other motions—possibly like these." The reports on the motions were taken after the second performance tests, about 8 to 10 minutes after the ride. At this time the subjects were first asked to describe the motions and any observed changes in them and then to identify them by means of any of the 25 diagrams in Figure 1 which corresponded to their perceptions. All curved paths shown on the diagrams could be interpreted by the subjects as representing either forward and backward movements or sideward movements. (Actually only 16 subjects reported any kind of sideward movement.)

## D. RESULTS

### 1. *Perceptual Shifts*

Of the 267 subjects, 93 felt only one kind of motion, 122 felt two kinds, 44 felt 3, six felt 4, and two felt 5 kinds.

Analysis of the sequences of perceptions showed little in the way of regu-

larity. The 410 recorded sequences were distributed among 51 different ones, counting all miscellaneous perceptions as a single kind of motion. Among the 122 sequences of two, 108 were couplings of Curve 1 (the actual path)

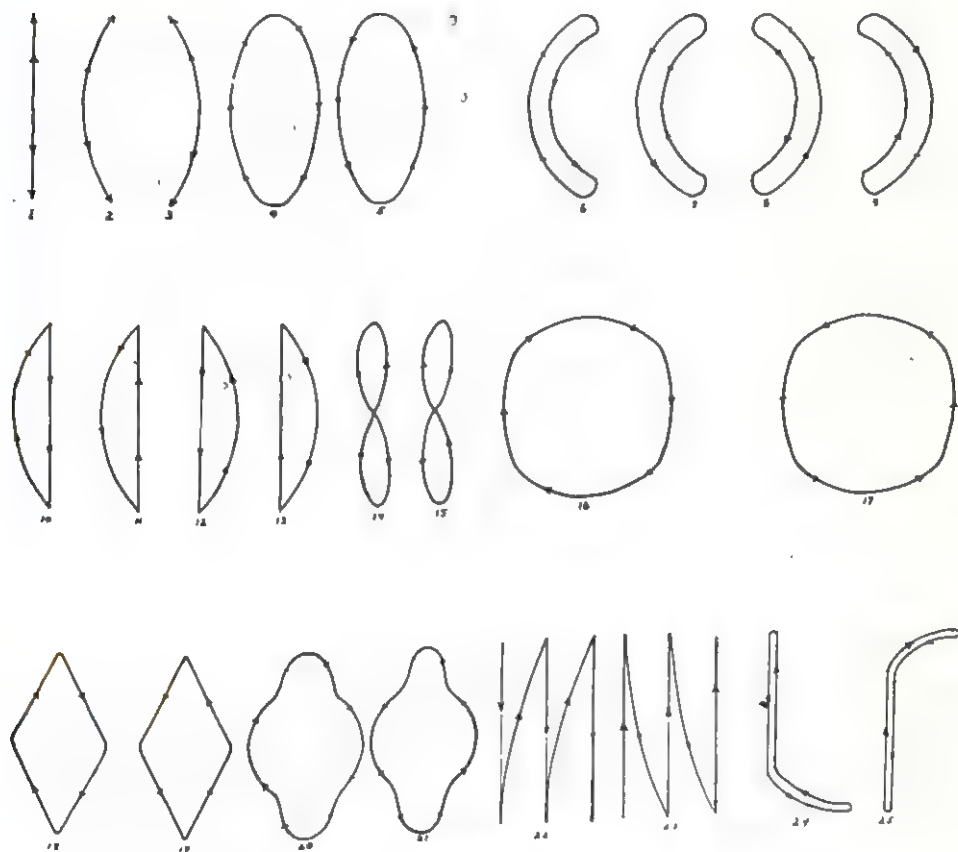


FIGURE 1

with another perception and 14 were pairs not involving 1. There were few other outstanding sequences. Curves 4 and 5 were reported by the same subject in 17 cases; Curves 4 and 10 or 5 and 11 in nine cases; Curves 4 and 12 in five cases.

## 2. Apparent Paths of Movement

Among those 93 subjects who reported only a single path of movement, 73 likened it to Curve 1, 10 to Curve 4, three to 10, three reported an intermittent downward movement, two to Curve 5, and one each to Curves 2 and 15.



The other 174 subjects reported more than one path of movement. All of the types of movement portrayed in Figure 1 were reported except Curves 9, 16, 17, 18, 19, 20, and 21. Curves 2, 3, and 8 were reported by less than one per cent of subjects; Curves 6, 14, 15, 22, 23, and 25, by one or two per cent; and the remainder by more. Table 1 shows the frequency of report of those curves named by more than 2 per cent of the subjects. It will be noted that while simple vertical motion (Curve 1) is the most fre-

TABLE 1  
APPARENT PATHS OF MOVEMENT

Apparent movement See Figure 1	Per cent of subjects reporting N = 267
1	
4	88%
5	31%
10	15%
12	11%
11	7%
2	4%
13	3%
24	3%
Others not among diagrams	3%
	16%

quent perception, the ellipse (Curves 4 or 5) is next most frequent, and then come variations of the ellipse (Curves 10, 11, 12, 13, and 24). Non-elliptical paths (2 and "others") are relatively infrequent. The category "others" refers to paths described by the subjects but not portrayed by the curves. These included perceptions as of swinging forward and backward, intermittent downward movements, forward progression as over waves, and others. The elliptical or near-elliptical patterns are interesting from the point of view of the direction of travel: in 128 cases the subject reported that he seemed to move forward as he began to come down (Curves 4, 6, 10, 13, and 23), while there were only 75 cases in which the elliptical movement went down-and-backward (Curves 5, 7, 8, 11, 12, and 22). Only 16 reported sideward movements of any kind.

### 3. Relation to Wave-Form

There was only one interesting relationship of perception to wave-form. In the *K*-wave, an asymmetrical wave with 0.65 *g* downward acceleration and 0.17 *g* upward acceleration, 6 out of 33 subjects did not report the upward acceleration at all, describing the wave as an intermittent downward movement. This sometimes occurred with other waves.

#### 4. Relation to Development of Sickness

The subjects were divided into the not-sick ( $N = 199$ ), the slightly sick ( $N = 40$ ), and the vomiters' ( $N = 28$ ). While there were no significant differences in the number of motions reported, the sick subjects showed a tendency to report a path other than the straight up-and-down as the predominant experience. Whereas 66 per cent of the not-sick named Curve 1 as the predominant one, only 45 per cent of the slightly sick and 43 per cent of the vomiters did so. In other words, sickness and perception of predominant non-linear paths of motion may possibly be associated.

#### E. DISCUSSION

It is not surprising to find variations in subjective reports under continuous repetitive stimulation. Such phenomena are familiar in the fields of visual and auditory perception. It is difficult, however, in our present state of knowledge of the receptors stimulated on the vertical accelerator, to formulate plausible hypotheses to explain the shifts of perception and to explain the differences between subjects. It may be that they depend in part on small variations or differences in head position. It is equally difficult to speculate on the slight relationship of the perceptions to sickness. Differences in perceived path of motion may be a causal factor in sickness, but it seems more likely that the perceptions were influenced by differences in attitude of susceptible and non-susceptible subjects and were thus less cause than effect.

#### REFERENCES

1. ALEXANDER, S. J., COTZIN, M., HILL, C. J., RICCIUTI, E. A., & WENDT, G. R. Wesleyan University studies of motion sickness: I. The effects of variation of time intervals between accelerations upon sickness rates. *J. of Psychol.*, 1945, 19, 49-62.
2. ———, Wesleyan University studies of motion sickness: IV. The effects of waves containing two acceleration levels upon sickness. *J. of Psychol.*, 1945, 20, 9-18.
3. ALEXANDER, S. J., COTZIN, M., KLEE, J. B., & WENDT, G. R. Studies of motion sickness: XVI. The effects upon sickness rates of waves of various frequencies but identical acceleration. *J. Exper. Psychol.*, 1947, 37, 440-448.

Department of Psychology  
University of Rochester  
Rochester 3, New York



## STUDIES OF MOTION SICKNESS: XVII. THE EFFECTS OF TEMPERATURE, POSTURE, AND WAVE FREQUENCY UPON SICKNESS RATES\*<sup>1</sup>

*Department of Psychology, University of Rochester*

CARSON JOHNSON AND G. R. WENDT

### A. INTRODUCTION AND SUMMARY

This study was designed to find how motion sickness rates on the vertical accelerator were affected by two temperatures, by four combinations of head and body posture, by five different waves and by three classes of susceptibility. These 14 variables were counterbalanced, others were approximately counterbalanced or were controlled. The subjects were blindfolded and subjected to motion for 20 minutes (or less if vomiting occurred). The results obtained on 120 men with amount of sickness represented by a *sickness index* giving single weight to nausea and double weight to vomiting

$$\frac{(f_{\text{nausea}} + 2f_{\text{vomiting}} \times 100)}{\text{No. of cases}}$$
 were: (a) There was no difference of

consequence between sickness rates at warm and cool temperatures: 86°F = 68, 65°F = 61. (b) One of the postures produced nearly twice the sickness of the others. The *sickness indices* were: seated with head upright, 47; seated with head back, 103; supine with head upright, 57; supine with head back, 53. (c) Analysis of the interrelations of the effects of body position, temperature, and type of wave seems to favor the conclusion that a simple mechanical action of wave motion on blood distribution was not a factor in the production of sickness. (d) The relation of *sickness index* to susceptibility was in the expected direction: susceptible, 100; intermediate, 65; non-susceptible, 30. (e) The relation of wave character to sickness rates was found to confirm the previous studies and yield one new result. Five acceleration levels were used, from 0.20 *g* to 0.65 *g*. The total energy per wave was kept at a constant value, a maximum mid-wave velocity of 400 ft./min. By this method wave frequencies of 13, 16, 22, 26, and 32 cycles/min. were pro-

\*Received in the Editorial Office on January 9, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This work has been supported by funds from a contract between the Office of Naval Research and the University of Rochester under contract N6ori-126, Task Order I.



duced. Corresponding *sickness indices* were 54, 83, 92, 63 and 34. The 26-cycle wave had not been previously used and was found to be above the optimum frequency (22 cycles) for production of motion sickness.

## B. PROCEDURES

The general procedures and apparatus of this study were identical (with certain exceptions noted below) with those of previous studies. [Former studies were done with a vertical accelerator at Wesleyan University; the present one with a similar apparatus at the University of Rochester (7). The chief differences between the Wesleyan and the Rochester vertical accelerator were that the sound level during operation had been very greatly reduced in the new machine, and that an effective two-way inter-com system was used.] Ninety male students of the School of Medicine and 30 college students in the Naval Reserve Officers' Training Corps were used. The experiment was set up to counterbalance 14 variables (5 types of wave, 4 postures, 2 temperatures, 3 susceptibility levels), approximately to counterbalance others (time of day, day of experiment, source of subject), and to control all other conditions.

### 1. *Counterbalanced Variables*

One subject was used in each of the 120 possible combinations of the 14 variables. The variables were as follows:

a. *Waves*. The vertical accelerator is like an elevator cab in which the controls are reversed each time the cab reaches its full upward or full downward velocity, so that it is in continuous up and down motion. The five waves were constant acceleration waves of which four were identical with those used in previous studies (see 5, p. 445 for references) and one was new. Each wave type had the same total energy per wave, having a mid-wave velocity of 400 feet per minute. By appropriate settings of the acceleration from 0.20*g* to 0.65*g* wave frequencies of 13, 16, 22, 26, and 32 cycles per minute were obtained. We refer to these by the code letters J'', H'', G', S, and A''. This is the third experiment in which A, H, and J waves have been used, the second for the G wave and the first for the S wave. The S wave was intermediate in frequency (26 cycles) between the most nauseating and least nauseating waves, 22 and 32 cycles. Physical characteristics of the waves are described in earlier reports (1, 2).

b. *Postures*. Each subject was either seated or supine, with head upright or head back. The seated subject used a cushion and a heavy lap-belt; the supine subject used a mattress, a lap-belt buckled across the hips, one below

the knees, and a foot-brace. Head posture was controlled by resting the back of the head against a solid, padded, adjustable rest, except that the "supine-head back" subjects used a pillow. The ear-eye line was marked by black eyebrow pencil. For upright head postures this line was kept horizontal; for backward postures it was kept vertical. Lapses from correct posture were corrected by the subject on instructions from the experimenter. In the supine posture with head up the shoulders were slightly elevated by an adjustable wooden incline. In each of the postures except the "supine-head back" there was slight bobbing of the head from wave action. This was greater with the faster waves.

*c. Temperatures.* Air was continuously recirculated through the cab. Within the circulating duct were electric heaters and a refrigerating coil. Air temperature of the cab was kept either at 65° F or at 86° F by thermostatic control. Each subject had come from out-of-doors and had then spent about 20 minutes in a room at about 74° F before going on the machine. All subjects removed coats and loosened collars.

*d. Susceptibility.* An equal number of subjects of each of three susceptibilities was used, susceptibles, intermediates, and non-susceptibles, as determined by scores on motion sickness history questionnaires (4). The dividing scores were so chosen as to allocate about one-third of the total population to each group (38-36, 35-30, 29 or less). Among the NROTC students all men had been on at least one ocean cruise. Only physically fit men were used. In this study, as in those previously reported, the selection of subjects could be substantially random with respect to other factors, since all students of the School of Medicine, except one, and all NROTC students volunteered.

As noted above there was approximate counterbalancing of certain other variables with respect to each of the 14 completely counterbalanced variables. Subjects were used at hours from 8:30 a.m. to 8:30 p.m. with an attempt to counter-balance time since last meal even though previous studies have shown that time of day is probably not a significant factor (3, 5). Similarly, the variables were distributed throughout the experiment so as to get day to day counterbalancing and were distributed equally among the medical and naval students.

## 2. Controlled Conditions

In this series of motion sickness experiments a great deal of effort has gone into control of the attitude of the subjects. In order to avoid distortions by rumor no advance notice was given to the subjects of the intention to ask for volunteers. The request for volunteers was orally presented by a standard appeal. This was followed by a written restatement to the subjects of the

facts orally presented. The motives appealed to in order to get volunteers were: the desire to aid experiments which were not being done elsewhere, the desire to be a 100 per cent volunteer group, an opportunity to learn something about motion sickness, willingness to take the "calculated risk" that one would fall in the five per cent selected who would vomit, and knowledge that each would serve but once and at his convenience. Each subject was also given two dollars. This was represented as being given, "not as pay for services, but to avoid hardship on the part of those who give up time from remunerative work." The five per cent "calculated risk" of becoming sick was accurately represented for the group as a whole, but not for any particular combination of variables.

Control of attitude at the time of the experiment, so as to duplicate that produced by the initial approach, was also given close attention. The same experimenter (C. J.) did all of the experimentation and used the same conversational controls and procedures with each subject.

Each subject went through the following procedure: (a) on arrival at the laboratory there were about two to three minutes of orienting conversation, carefully planned to have the experimenter retain the lead; (b) the subject went through a practice session with the Johnson-Paschal code substitution test; (c) he was prepared for the experiment by drawing the ear-eye line, adjusting the equipment for control of posture, and giving instructions; (d) a pre-exposure trial on the code substitution test was given; (e) the subject was again placed in the cab, blindfolded with rubber goggles with blackened glasses, and exposed to a particular combination of conditions for 20 minutes or until he appeared to begin to vomit; (f) after stopping the motion he rested in the cab for two minutes; (g) he again took the code substitution test; (h) a questioning session and conference lasting about 20 minutes filled the remainder of the hour. Data on nausea and other subjective effects (including the apparent path and magnitude of the motion) were obtained. The data on performance tests, on the apparent magnitude and path of motion, on the time taken to produce vomiting, and other observations will be reported at a later date.

The data here reported were obtained between December 13, 1949, and May 1, 1950.

### 3. *Measures of Sickness*

The indicators of sickness were the same as in the earlier studies in this series except that sweating associated with nausea was not used because of a lack of data on its relation to air temperature. The procedure of this study also differed from the Wesleyan studies in that the subjects were asked to

report nausea as soon as they felt it, rather than waiting until after the exposure. Two indicators of sickness were used, unequivocal report of nausea and retching or vomiting. Agreement of two out of three judges who read each protocol was required to classify a subject. Three criteria for evaluation of the effects of the experimental conditions were used: *Criterion 1* = frequency of retching or vomiting, *Criterion 2* = frequency of retching or vomiting, plus frequency of nausea alone, *Criterion 3* = retching or vomiting weighted 2, plus nausea alone weighted 1. In order to allow direct comparisons of *Criterion 3* between conditions or experiments having differing numbers of subjects *Criterion 3* was also converted to the *Sickness Index*,

$$\text{Criterion 3} \times 100$$

\_\_\_\_\_  
No. of cases

Data on time taken to produce nausea, time taken to produce vomiting and on duration of sickness were also obtained.

### C. RESULTS AND DISCUSSION

The data of this study are displayed in three tables. Table 1 presents the raw data for each of the 120 subjects used in the 120 combinations of variables. These are presented to enable others to make computations omitted by us. Table 2 shows sickness rates at each posture, temperature, wave frequency, and susceptibility. Table 3 displays a break-down of results for certain combinations of variables selected to indicate the interaction of the experimental conditions and to aid in estimates of the correctness of the results. Certain other data not displayed in tables are presented in the text.

In Table 1, which gives the raw data, each row shows the data of four different subjects: serial number of subject, time to onset of nausea, and time to onset of vomiting (or retching). Consecutive rows are arranged so as to separate subjects by susceptibility category, cab temperature, and wave frequency. The four groups of subject columns are arranged so as to separate the subjects according to the four postures used. Where a— appears in the table under nausea or vomiting, it means that that subject did not report nausea or did not vomit. Where a+ appears, the subject reported during the interview that he had been nauseated during the exposure to motion. Such reports, since they were not volunteered during motion, may lie somewhere between a fully acceptable and an unacceptable evidence of unequivocal nausea. In the analysis of data they were counted as being unequivocal nausea because the subjects felt sure of their judgments.

Subjects numbered 1-90 were medical students; 91-120 were NROTC students.





TABLE 1 (continued)

Wave freq. cyc/ min	T °F	Susc.	Seated, head up Time to naus.	Seated, head up Time to vom.	Seated, head back Time to vom.	Seated, head up Time to naus.	Seated, head back Time to vom.	Subj. No.	Seated, head up Time to vom.	Seated, head back Time to vom.	Subj. No.	Seated, head up Time to naus.	Seated, head back Time to vom.	Subj. No.	Seated, head up Time to naus.	Seated, head back Time to vom.
	86	3	46	12' 10"	—	117	3' 40"	4' 55"	84	1' 30"	7' 00"	85	—	—	—	—
	65	1	43	—	—	34	—	—	7	—	—	68	+	—	—	—
	65	2	94	—	—	60	19' 40"	—	69	—	—	—	8' 00"	9' 30"	—	—
	65	3	107	14' 05"	14' 40"	78	7' 25"	19' 10"	31	19' 30"	—	120	—	—	—	—
26	86	1	51	—	—	105	—	—	75	—	—	27	—	—	—	—
	86	2	114	—	—	80	—	—	20	—	—	102	+	—	—	—
	86	3	38	13' 00"	13' 15"	116	19' 00"	—	58	+	—	62	+	—	—	—
	65	1	9	—	—	13	+	—	100	—	—	53	—	—	—	—
	65	2	76	—	—	26	8' 30"	14' 15"	28	—	—	4	+	—	—	—
	65	3	88	+	—	52	14' 30"	—	119	15' 25"	15' 30"	79	5' 30"	—	—	—
32	86	1	30	—	—	83	—	—	55	—	—	109	—	—	—	—
	86	2	87	—	—	36	16' 00"	16' 20"	113	—	—	81	—	—	—	—
	86	3	11	—	—	63	—	—	39	—	—	42	—	—	—	—
	65	1	111	—	—	108	—	—	72	—	—	25	—	—	—	—
	65	2	37	12' 00"	—	3	—	—	14	—	—	96	9' 30"	—	—	—
	65	3	56	—	—	44	3' 50"	13' 10"	97	+	—	54	+	—	—	—

TABLE 2  
EFFECTS OF TEMPERATURE, POSTURE, SUSCEPTIBILITY, AND WAVE TYPE ON  
MOTION SICKNESS  
(120 S's)

Variable	Nausea		Criterion 1 Vomiting		Criterion 2 Vomiting plus nausea		Criterion 3 Vomiting = 2 Nausea = 1		Sickness Index
	N	%	N	%	N	%	Nausea = 1		
86° F	11	18	15	25	26	43	41		68
65° F	17	28	10	17	27	45	37		61
Seated, head up	4	13	5	17	9	30	14		47
Seated, head back	9	30	11	37	20	67	31		103
Supine, head up	5	16	6	20	11	37	17		57
Supine, head back	10	33	3	10	13	43	16		53
Susceptible	16	40	12	30	28	70	40		100
Intermediate	8	20	9	23	17	43	26		65
Non-susceptible	4	10	4	10	8	20	12		30
J", 13 cycles	3	12	5	21	8	33	13		54
H", 16 cycles	8	33	6	25	14	58	20		83
G', 22 cycles	4	17	9	38	13	54	22		92
S, 26 cycles	9	38	3	13	12	50	15		63
A", 32 cycles	4	17	2	8	6	25	8		34

Table 2, which shows the effects upon sickness rates of each of the 14 main variables of the experiment, indicates the following obtained results: (a) *Susceptibility of subject*. The results were like those of all earlier studies (4, 5), showing that previous history of motion sickness gave a fair prediction of sickness on the vertical accelerator. (b) *Wave frequency*. The results were like those of four earlier studies (for summary see 5, p. 445) in showing increase of sickness with increase of wave frequency up to 22 cycles per minute, and then a sharp decrease to a low sickness rate at 32 cycles. The 26-cycle wave, not previously used, produced a sickness rate intermediate between the very high rate of the 22 cycle wave and the very low rate of the 32 cycle wave. (c) *Temperature*. The warm cab (86° F) was slightly more effective in producing sickness than the cool cab (65° F) according to *Criteria 1 and 3*. The differences were too small to be considered of consequence. (d) *Posture*. The results were that the "seated with head back" posture produced nearly twice as much sickness as each of the other three postures which were about equal to one another. It is possible that the "supine with head back" posture was actually more favorable for avoidance of sickness than indicated. Six of the eight cases of nausea on which one judge disagreed with the other two were in this posture. If these were eliminated the *sickness index* for that posture would be 33 instead of 53. According to previous work the "seated with head back" posture should have been very favorable, as should the "supine with head back" posture (6, p. 326-9). It is possible that in our experi-

ment the former head posture was awkward, producing traction on the throat during motion.

Table 3 shows the inter-relations of the fourteen variables in terms of the *sickness index*. The purpose of this table is to allow one to make estimates of the trustworthiness of the results obtained in this experiment from their correspondence or lack of correspondence with previous experiments. From this one sees that while in general the results of the known variables (susceptibility and four of the waves) show the expected trends in each fraction of the data of the unknown variables, there are a number of exceptions, especially in small fractions of the data.

TABLE 3  
INTER-RELATIONS OF VARIABLES ACCORDING TO SICKNESS INDEX  
(Figures comparable in magnitude but of unequal reliability\*)

Variables	Seated, head up	Seated, head back	Supine, head up	Supine, head back	Suscep- tible	Inter- mediate	Non-sus- ceptible	86°	65°
<i>Waves</i>									
13 cyc/min	67	83	33	33	113	0	50	67	42
16 cyc/min	33	167	83	50	88	113	50	108	58
22 cyc/min	83	117	83	83	125	113	38	108	75
26 cyc/min	50	83	50	67	100	50	13	42	75
32 cyc/min	17	67	17	33	50	50	0	17	50
<i>Temperatures</i>									
86° F	60	93	73	53	100	71	31	—	—
65° F	40	113	33	53	100	55	21	—	—
<i>Susceptibility</i>									
<i>Categories</i>									
Susc.	110	130	100	60	—	—	—	—	—
Intermed.	20	110	40	90	—	—	—	—	—
Non-susc.	20	70	20	10	—	—	—	—	—

\*Numbers of cases in each cell are as follows: Wave *vs* posture, N = 6; wave *vs* susceptibility, N = 8; wave *vs* temperature, N = 12; temperature *vs* posture, N = 15; temperature *vs* susceptibility, N = 20; susceptibility *vs* posture, N = 10.

By inspection of Table 3 one can estimate that the data on relation of sickness to posture show the following order of consistency with expectations based on previous studies: first, and most likely to be correct, "supine, head up," second, "seated, head back," third, "seated, head up," last, "supine, head back." This tends to give confidence in the relatively high sickness rate obtained in the "seated, head back" posture, but to cast serious doubt on the data of the "supine, head back" posture. One cannot say from these considerations whether the obtained figure in the latter posture is too high or too low. One may note by reference to the raw data in Table 1 that six of the subjects on whom only two of three judges agreed were in this posture. Inspection of the interrelations of temperature with wave-frequency and with



susceptibility shows fairly good internal consistency with expectation based on previous studies, suggesting that the obtained negative results in relation to temperature may be correct. Inter-relations of wave-frequency and susceptibility at the 22 cycle and 26 cycle waves suggest that the finding of less sickness at 26 cycles is a trustworthy result.

Other calculations have been done in an attempt to test certain hypotheses about the causes of sickness. These were made from the raw data of Table 1, but will not be displayed in tables. Using *Criterion 3* (nausea = 1, plus, vomiting = 2), ratios of sickness were calculated for a number of combinations of variables. To test for the hypothesis that movement of the viscera caused sickness it was assumed that they would move more in the seated than in the supine position and more with rough waves than with slow waves. Ratios of *Criterion 3*, seated to supine at each wave frequency, show that there was more sickness for seated subjects at all frequencies but no relationship of the seated/supine ratio to frequency. The hypothesis that movement of the viscera caused sickness is therefore not supported. To test for the hypothesis that alterations in blood distribution under the influence of wave action caused sickness it was assumed that such alterations would be greater in the seated than in the supine position, greater with rough waves than with slow waves, and greater at 86° F than at 65° F (because of vasodilation). Ratios of *Criterion 3*, warm/cool at each wave frequency and seated-warm/seated-cool at each wave frequency, rather than confirming this hypothesis, appear to show the opposite. The obtained results were that 86° F produces less sickness than 65° F with the 32 cycle and 26 cycle waves and more sickness with the 22, 16, and 13 cycle waves and that this superiority of a warm cab with the rough waves was as great in the supine postures as in the seated postures.

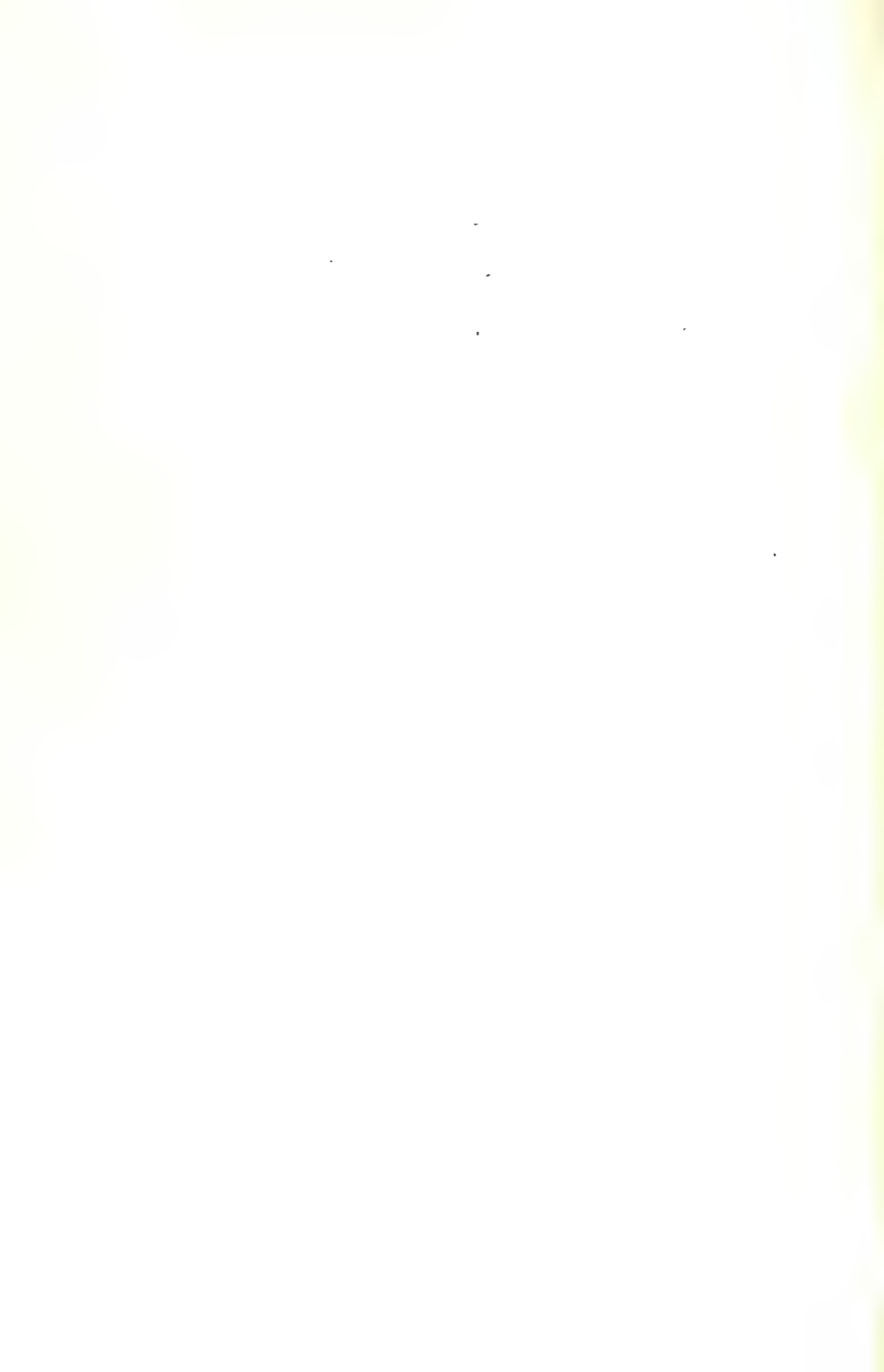
Another analysis showed that the medical students were more susceptible than the NROTC students.

#### REFERENCES

1. ALEXANDER, S. J., COTZIN, M., HILL, C. J., RICCIUTI, E. A., & WENDT, G. R. Wesleyan University studies of motion sickness: I. The effects of variation of time intervals between accelerations upon sickness rates. *J. of Psychol.*, 1945, **19**, 49-62.
2. ———: III. The effects of various accelerations upon sickness rates. *J. of Psychol.*, 1945, **20**, 3-8.
3. ———: V. Incidence of sickness at various hours of the day. *J. of Psychol.*, 1945, **20**, 19-24.
4. ———: VI. Prediction of sickness on a vertical accelerator by means of a motion sickness history questionnaire. *J. of Psychol.*, 1945, **20**, 25-30.

5. ALEXANDER, S. J., COTZIN, M., KLEE, J. B., & WENDT, G. R. Studies of motion sickness: XVI. The effects upon sickness rates of waves of various frequencies but identical accelerations *J. Exper. Psychol.*, 1947, **37**, 440-448.
6. TYLER, D. B., & BARD, P. Motion sickness. *Physiol. Rev.*, 1949, **29**, 311-369.
7. WENDT, G. R. Committee on Aviation Medicine, Final Report. National Research Council, 1945.

*Department of Psychology*  
*University of Rochester*  
*Rochester 3, New York*



## LYSERGIC ACID DIETHYLAMIDE (LSD-25): V. EFFECT ON SPATIAL RELATIONS ABILITIES\*

*Departments of Medicine, Neurology, and Psychiatry, The Mount Sinai Hospital,  
New York City*

H. A. ARRAMSON, M. E. JARVIK, M. W. HIRSCH, AND A. T. EWALD<sup>1</sup>

### A. INTRODUCTION

In view of the reports that subjects under the influence of lysergic acid diethylamide (LSD-25) frequently manifest incoherence and confusion (7, 8) and often report visual and other sensory misperceptions and hallucinations (1, 10) it seemed desirable to determine whether LSD-25 produced impairment in a subject's abilities to perform adequately on tests of spatial relations. It is the purpose of this paper to investigate the effects of lysergic acid diethylamide on scores obtained on two tests of spatial orientation, the Thurstone Hand Test and the Revised Minnesota Paper Form Board Test (MPFB).

### B. METHOD

#### 1. Subjects

The population of this study consisted of 12 non-psychotic, adult volunteers, whose median age was 27, ranging from 21 to 33. Six of the subjects were male and six female. Preliminary psychiatric interviews and clinical psychological tests were administered to the subjects for purposes of screening. Only those who were considered non-psychotic were accepted. All subjects were of average intelligence or above and eight were of superior intelligence, according to the Wechsler-Bellevue Intelligence Scale. Seven were graduate students in psychology or allied fields. There were two housewives, a hairdresser, an auditor, and a public relations worker.

#### 2. Tests

Two tests were used in this study. The first was the Thurstone Hand Test devised by L. L. Thurstone in 1925 (9). It consisted of 49 drawings of right and left hands in various positions. The subject had to indicate

\*Received in the Editorial Office on January 10, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was aided by a grant from the Geschickter Fund for Medical Research. We are indebted to Sandoz Pharmaceuticals, Incorporated, for supplies of LSD-25 and other compounds.



whether each drawing represented the right or left hand. The same form of the test was used for each testing session.

The second test of spatial relations was the Minnesota Paper Form Board Test prepared by Likert and Quasha (6). It consisted of problems in which a geometrical figure was broken into parts. The subject had to select from five complete geometric figures the one which represented the parts correctly fitted together. Two forms of this test were available; each was divided into two comparable forms by considering the odd numbered problems within each form as one form, and the even numbered problems as a second form.

### 3. Procedure

*a. Administration of drug.* At approximately 9:30 A.M., subjects were given the drug orally in 75 cc of water. The drug administration and testing took place in an air-conditioned room where groups of three, four, or five subjects were tested together. All the subjects had previously been in this room and had met the experimenters during preliminary interviews and the clinical psychological tests. They may or may not have been acquainted with other subjects in the group. Two experimenters administered the tests. The drug doses varied among the subjects tested in any one day.

Each subject was tested three times, and eight of the subjects were tested four times, with at least one week between experiments. The first time, the subjects received a zero dose (placebo consisting of 75 cc of tap water). Because LSD-25 is a tasteless, odorless, and colorless substance, the water dose served as a control for the unknown variables in the test situation. Subjects were never told what their drinks contained. A dose of 50 micrograms of LSD-25 was given the second time subjects were tested, and 100 micrograms the third time. Those subjects available for a fourth testing session again received a zero dose of the drug. In view of the nature of the drug it seemed inadvisable to administer a high dose of the drug the first day. By varying the doses used in any one day and using a given test form in one day, all forms were used at each dose level of testing. The second zero testing was administered as a means of evaluating the effect of repeated testing with the same tests.

Subjects receiving the drug had eaten no food since the previous evening. One-half hour after the drug or placebo was administered, a light breakfast was given to the subject. No stimulants were permitted and no smoking was allowed during the experimental period.

*b. Administration of tests.* The two tests were administered approximately  $1\frac{3}{4}$  hours after the subjects had received the drug. The Thurstone

Hand Test was administered first, and was followed by the Minnesota Paper Form Board Test. These two tests were given to subjects immediately after a series of nine tests of immediate memory (4), eight cancellation tests (3), and two forms of an arithmetic test (5). The latter tests were administered to the subjects between one hour and  $1\frac{3}{4}$  hours after the ingestion of the drug.

The subjects were instructed in the Thurstone Hand Test to indicate whether the pictures represented left or right hands by checking the corresponding box below the picture. They were allowed one minute to do as many of the problems as they could and then immediately repeated the test. The first test was called Form I, the second Form II.

In the Minnesota Paper Form Board Test the subjects selected the one of five composite pictures which to them represented the separate scrambled parts properly assembled. They first did the practice problems and were checked to be sure they had done them correctly. They were then allowed five minutes to do as many problems as they could. A different form was administered at each testing session.

*c. Method of analysis.* The Thurstone Hand Test was scored by counting the number of correct responses and subtracting the number of incorrect responses from that number as a correction for guessing. The Minnesota Paper Form Board Test was scored similarly. Since there were five possible choices, the score was the number of correct responses minus  $1/5$  of the number of incorrect responses.

Each test was then analyzed in the same way. The scores under zero micrograms were compared with those obtained under 50 micrograms and under 100 micrograms, respectively, by means of Wilcoxon's non-parametric test of paired replicates (11), and the Spearman rank-order correlation (2). With these methods, the significance of the difference and the relationship between scores under two different doses were obtained. The 50 microgram and 100 microgram scores were compared in the same way.

Wilcoxon's non-parametric test of paired replicates was also applied to the data from the first zero testing (pre) and the second zero testing (post) to determine the significance of the difference in scores, and therefore, to determine whether repetition was effective in changing the scores.

### C. RESULTS

The results are summarized in Figures 1 and 2 and Tables 1 and 2. Figure 1 demonstrates the changes which occur in the test performance under increasing doses of LSD-25. The average scores on the Minnesota Paper

Form Board Test and on the two forms of the Thurstone Hand Test are plotted for the pre-drug zero, the 50 microgram, and the 100 microgram sessions. In both forms of the Thurstone Hand Test there was a negative

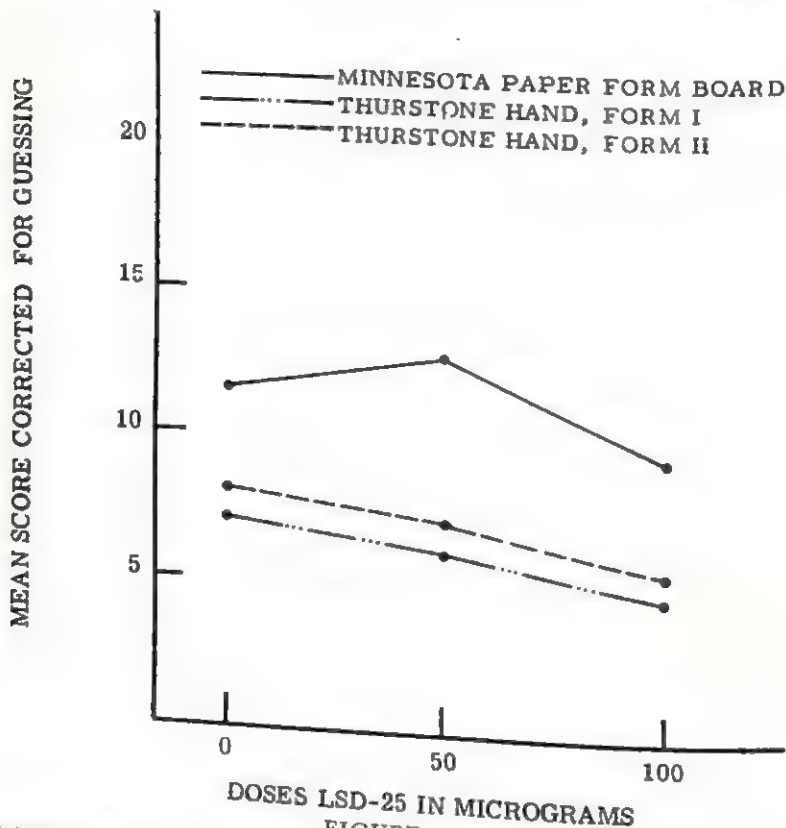


FIGURE 1

Mean scores, corrected for guessing, on the Minnesota Paper Form Board Test and on two identical forms of the Thurstone Hand Test, for nine and twelve subjects, respectively. Subjects were tested under zero and two LSD-25 doses.

correlation between score and drug; that is to say, the higher the dose of LSD-25 the lower the average score on the tests, for twelve subjects.

The mean score on the Minnesota Paper Form Board Test increased under 50 micrograms as compared to the pre-drug zero. Under 100 micrograms, the mean score was lower than under either of the two other doses. Here the *N* was 9.

The significance of the differences between scores under two different doses appears in Table 1. Four comparisons were made: pre-drug zero with 50 micrograms, pre-drug zero with 100 micrograms, 50 micrograms with 100

micrograms, and pre-drug zero with post-drug zero. The latter will be noted subsequently. On Form I of the Thurstone Hand Test the only score decrease which was significant (at the .05 level) was that between zero and

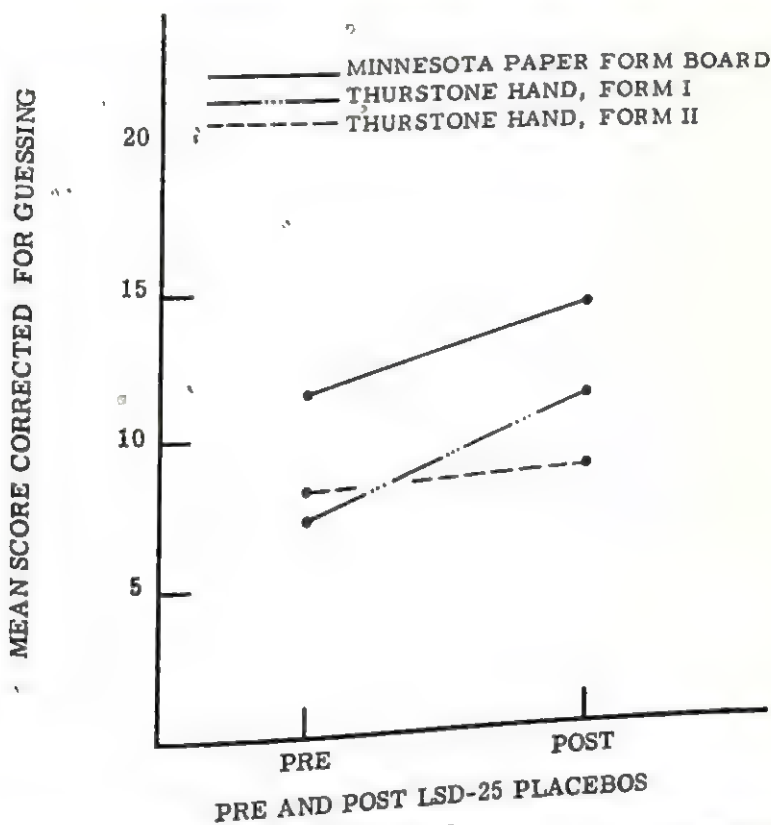


FIGURE 2

Mean scores, corrected for guessing, on the Minnesota Paper Form Board Test and on two identical forms of the Thurstone Hand Test, for six and eight subjects respectively. Subjects were tested under a zero dose before and after the drug tests.

100 micrograms. The drop between zero and 100 micrograms was also the only significant change on Form II. The Minnesota Paper Form Board Test showed a significant lowering of score from 50 to 100 micrograms, only. All other differences among the first three dose groups could have occurred by chance more than 5 times in 100.

The comparisons between the pre-drug zero and the post-drug zero sessions yielded results summarized in Figure 2 (and Table 1). The mean scores of pre- and post-drug tests are plotted for both forms of the Thurstone Hand Test and for the MPFB test. All scores, on the average, in-



creased in the post-drug tests. The improvement on Form I of the Thurstone Hand Test was significant at better than the .02 level. The improvement on the MPFB test and on Form II of the Thurstone Hand Test were chance occurrences. The *N* was 8 on the Thurstone Hand Test and 6 on the MPFB test.

TABLE 1  
PROBABILITY OF DIFFERENCES BETWEEN DOSES OCCURRING BY CHANCE  
(Minnesota Paper Form Board and Thurstone Hand Test)  
*N* = 12

Test	Probability of score differences occurring by chance			
	Doses Zero and 50	LSD-25 Zero and 100	compared, in micrograms 50 and 100	Zero-Pre and Zero-Post
Thurstone Hand I	>.05	.05	>.05	<.02**
II	>.05	<.05	>.05	>.05**
Minn. Paper Form Board	>.05*	>.05*	.02	>.05***

\**N* = 9.

\*\**N* = 8.

\*\*\**N* = 6.

TABLE 2  
SPEARMAN RANK-ORDER CORRELATION COEFFICIENTS RELATING SCORES UNDER DIFFERENT DRUG DOSES  
(Minnesota Paper Form Board and Thurstone Hand Test)

Test	Doses LSD-25 correlated, in micrograms			<i>N</i>
	Zero and 50 <i>p</i>	Zero and 100 <i>p</i>	50 and 100 <i>p</i>	
Thurstone Hand I	.21	.41	.41	12
II	.37	.46	.63*	12
Minn. Paper Form Board	.31	.52	.59**	9

Note: *p* equals the rank order correlation coefficient.

\*Significant at .05 level; all other correlations are attributed to chance.

\*\**N* = 11.

The Spearman rank-order correlation coefficients, relating scores under the various drug doses, appear in Table 2. Only one of the correlation coefficients was significant at the .05 level—that between the 50 and 100 microgram sessions on Form II of the Thurstone Hand Test. The coefficient was 0.63. All other correlations were lower and not significantly different from zero.

#### D. DISCUSSION

Little of the data obtained on the two tests of spatial relations indicates significant changes between non-drug and drug states of our subjects because

of the small number of subjects. However, it should be noted that scores under the drug were almost always lower than those under the placebos. The improvement on the post-drug placebo as compared to the pre-drug placebo suggests that successive administration of the tests would result in successively higher scores. It appears, therefore, that LSD-25 interferes with the normally anticipated improvement and, in addition, tends to reduce the original level of performance. The changes reported in the tables would probably be more significant if comparison were made with post-drug rather than pre-drug placebo scores and if the number of subjects were larger.

It is interesting that, despite the impaired performance under the drug, under the second placebo there was an improvement in performance. This suggests that learning may have taken place even under the influence of the drug.

The MPFB showed improvement from the pre-drug zero scores to the 50 microgram scores.<sup>4</sup> We do not know whether learning would have been even greater without LSD-25 or whether this drug dose has no effect upon the MPFB test; under 100 micrograms there was impairment of performance.

Likert and Quasha (6) report that the MPFB test is moderately related to tests of intelligence and of mechanical ability. Consequently, impairment of ability on the MPFB under the influence of LSD-25 suggests that this drug also interferes somewhat with intellectual and mechanical functioning.

All the rank-order correlations but one could have occurred by chance with this particular number of subjects. This suggests that knowledge of the subject's performance under a first control is of little value in predicting his rank position under the drug. However, all correlations were positive suggesting some degree of consistency in response under different doses. The highest correlations occurred between the two drug doses. It seems easier to predict a subject's test performance under a given dose of LSD-25 from his performance under another LSD-25 dose, than from his test performance under a placebo.

#### E. SUMMARY AND CONCLUSIONS

Twelve non-psychotic adults were tested with the Thurstone Hand Test and the Minnesota Paper Form Board Test under 50 and 100 microgram doses of lysergic acid diethylamide, and under zero doses, both before and after the LSD-25 tests. The following conclusions can be made:

1. LSD-25 tends to impair the subject's scores on tests of spatial rela-

tions, but with the *N* used, only some of the changes were significant. On both forms of the Thurstone Hand Test there was significant impairment only between zero and 100 microgram scores. On the MPFB the only significant change was between the two drug doses.

2. The tests showed that on subsequent testing without the drug, subjects performed more proficiently than at the first testing.

3. There was only one significant rank order correlation between scores under the various drug and placebo doses. However, all were positive, and those between the two drug doses were highest.

#### REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., LEVINE, A., KORNETSKY, C., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psychol.*, 1955, **39**, 3-60.
2. GARRETT, H. Statistics in Psychology and Education. New York: Longmans, Green, 1953. Pp. 451.
3. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): IV. Effect on attention and concentration. *J. of Psych.*, 1955, **39**, 373-383.
4. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): VI. Effect upon recall and recognition of various stimuli. *J. of Psych.*, 1955, **39**, 443-454.
5. JARVIK, M. E., ABRAMSON, H. A., HIRSCH, M. W., & EWALD, A. T. Lysergic acid diethylamide (LSD-25): VIII. Effect on arithmetic test performance. *J. of Psych.*, 1955, **39**, 465-473.
6. LIKERT, R., & QUASHA, W. Revised Minnesota Paper Form Board Test. New York: Psychological Corp., 1948. Pp. 16.
7. RINKEL, M., DESHON, H. J., HYDE, R. W., & SOLOMON, H. Experimental schizophrenia-like symptoms. *Am. J. Psychiat.* 1952, **108**, 572-578.
8. SAVAGE, C. Lysergic acid diethylamide (LSD-25): A clinical-psychological study. *Am. J. Psychiat.*, 1952, **108**, 896-900.
9. STOELTING, C. H. Manual: 27113—Thurstone Hand Test. Chicago: Stoelting, (mimeograph).
10. STOLL, W. A. D-lysergic acid diethylamide (LSD-25): A phantasticum of the ergot group. *Schweiz. Archiv. f. Neurolog. u. Psychiat.*, 1947, **60**, 1-45.
11. WILCOXON, F. Some Rapid Approximate Statistical Procedures. New York: American Cyanamid Co., 1949. Pp. 16.

133 East 58 Street  
New York City 22

## LYSERGIC ACID DIETHYLAMIDE (LSD-25): VI. EFFECT UPON RECALL AND RECOGNITION OF VARIOUS STIMULI\*

*Departments of Medicine, Neurology, and Psychiatry, The Mount Sinai Hospital,  
New York City*

M. E. JARVIK, H. A. ABRAMSON, AND M. W. HIRSCH<sup>1</sup>

### A. INTRODUCTION

Previous studies by the present authors (1) upon physiological effects of LSD-25 suggested that there would be some interference in the normal integrative processes of the nervous system which might affect the factors of acquisition and retention. Consequently, in this study, a series of memory tests of varying difficulty, some auditory and some visual, were presented to the same subjects.

Only two studies in the literature have attempted to objectively measure memory changes induced by LSD-25. Sloane (7) reports slight but significant decreases on the Wechsler-Bellevue Memory Scale in his group of "healthy controls," schizophrenics, and depressives. He found no significant decreases in the accuracy of recall of a name, address, and message five minutes after learning. He also reported that digit span showed no significant change in any group. Savage (6) reported that there was no impairment of memory as measured by the Wechsler-Bellevue Intelligence Scale. In this laboratory unpublished studies with the Wechsler-Bellevue Intelligence Scale show that under the drug there were significant changes in all memory tests with the single exception of the digit span test.

It is the purpose of this study to investigate the effects of two doses of LSD-25 and a placebo upon the recognition and recall of verbal and visual stimuli. The effects of practice in taking the tests will also be analyzed and the relationship between a subject's performance under a placebo and under the drug will be determined.

### B. METHOD

#### 1. Subjects

The subjects were 12 paid, adult volunteers, six of whom were male and six female. Only those who were considered non-psychotic on the basis of

\*Received in the Editorial Office on January 10, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.  
<sup>1</sup>This investigation was aided by a grant from the Geschickter Fund for Medical Research. We are indebted to Sandoz Pharmaceuticals, Incorporated, for supplies of LSD-25 and other compounds.



a psychiatric interview and who were believed non-psychotic according to a battery of clinical psychological tests (1) were accepted.

The median age was 27 years, ranging from 21 to 33. The group was at least of average intelligence and eight subjects were in the superior intelligence range, on the Wechsler-Bellevue Intelligence Scale. Seven subjects were graduate students in psychology or allied fields, two were housewives, one an auditor, one a hairdresser, and one a public relations worker.

## 2. *Apparatus and Tests*

*a. Variety.* Five visual and four auditory tests were used, containing stimuli of varying degrees of familiarity. Visual stimuli were projected on a screen approximately 6 to 10 feet from the subjects by means of an SVE slide projector with an Airequipt Slide Changer. Tests of this kind were object recall, word opposite recall, nonsense syllable recall, nonsense syllable recognition, and letter recognition. The auditory tests, which were presented verbally by the experimenter, included digit span, sentence recall, unrelated word recall, and paired associate recall.

*b. Source and description.* Several of the tests were selected from Wells and Ruesch Mental Examiner's Handbook (10), namely: objects, word opposites, and digit span. The object test contained 10 slides each with a line drawing of a simple common object. For word opposites, 10 words, all of which had common antonyms, were projected one at a time. The digit span consisted of a series of four to nine numbers.

Ten nonsense syllables of low association value, as prepared by Melton and Malmo (4) were used for nonsense syllable recall. Five of these together with five other syllables of low association value were arranged in random order to make up the nonsense syllable recognition test.

The letter recognition test consisted of a series of 10 letters, randomly selected from a page of print and presented on a single slide. No two letters were repeated. Ten additional slides contained one letter each, five of which were taken from the original 10 and were randomly arranged among five "new" letters.

Five couplets (in iambic pentameter) randomly selected from Shakespeare's *Romeo and Juliet* comprised the material for sentence recall. Thus, the subjects received 10 lines of 10 syllables each.

Ten unrelated words of comparable frequency of use in the English language were randomly selected from Thorndike (8). The paired associate test contained 10 pairs of words of varying degrees of relatedness and were taken from Melton and Safier (4).

Six equivalent but different forms of each test were prepared.

### 3. Procedure

*a. Administration of drug.* The drug was administered orally in 75 cc of tap water at approximately 9:30 a.m. Subjects ate no food since the previous evening; one-half hour after the ingestion of the drug they were given a light breakfast. No stimulants were permitted and smoking was not allowed during the testing day.

All 12 subjects were tested three times and eight participated in a fourth session. The dose given to all subjects at their first session was 75 cc of tap water, only. Since LSD-25 is a colorless, odorless, and tasteless substance, the water placebo served as a control for the test situation. Fifty micrograms of LSD-25 were given at the second session and 100 micrograms of LSD-25 at the third. For the fourth session subjects were again given a placebo of tap water. There was an interlude of at least one week between sessions.

Practice, anxiety, and other factors which might influence the test results were not controlled by the more usual experimental design of balanced order. The action of the drug upon a given individual is rather variable and not yet predictable, so it seemed inadvisable to begin with a high dose. The scores obtained under the second placebo served as indices of the effects of practice.

Both the drug administration and testing took place in an air-conditioned room where the subjects were tested in groups of 3, 4, or 5. The doses in any one day usually varied among subjects. The subjects had already been in the room and had met the experimenters during their preliminary interviews and psychological testing. They were sometimes acquainted with the other members of the experimental group.

*b. Administration of tests.* It was shown that LSD-25 begins to take effect about one-half hour after ingestion and evokes a maximum effect between one and 2½ hours, depending on the dose (1). Thus, one hour after ingestion of the drug the series of memory tests was administered for approximately 15 minutes. Visual tests were presented first. Slides were projected for one second with one second pause between slides except for the recognition tests when the pause was long enough for the subject to write a "yes" or "no" response. Following the 10 slides of the recall tests the subjects had to write down as many items as they could.

The auditory tests followed. Stimuli were read in a monotone by the experimenter at the rate of one per second. Sentences were read in couplets, and the subjects then wrote what they recalled. After each series of digits, and after the 10 unrelated words subjects recalled and wrote whatever they

remembered. After learning the series of 10 paired words, the subjects were given the 10 words randomly rearranged, one at a time, and after each had to try to write the word originally paired with it.

Each time the subject was tested, a different form of the test was used. In a given testing day only one form was administered.

*c. Analysis of data.* The Wilcoxon non-parametric test of paired replicates (9) was used to test the hypothesis that the group under one dose of LSD-25 responds differently than under another dose. Test scores on the recall tests were the number of correct responses; on the recognition test, scores were the number of correct minus the number of incorrect responses. Comparisons were made between the placebo and 50 microgram, between the placebo and 100 microgram, and between the 50 and 100 microgram sessions.

The same method was used to compare the two placebo testings to determine whether there was a difference in performance. Unlike the familiar *t* test the use of this technique does not require the assumption of a normally distributed population. Relationship between performance under two doses was determined by a Spearman rank-order correlation coefficient (3).

### C. RESULTS

The results are shown in Figures 1 through 4 and Tables 1 and 2. The mean scores obtained by the group of 12 on the visual memory tests under three doses of LSD-25 are given in Figure 1. For the most part, the scores decreased from an initially high value at zero micrograms to a lower value at 50 micrograms of LSD-25, and to a still lower value at 100 micrograms. Exceptions were the word opposite test where there was a slight increase from the zero to the 50 microgram score, and the nonsense syllable recognition where there was no score change between the 50 and 100 microgram level. In general, the highest scores were obtained on the recognition tests, and the lowest on the recall of nonsense syllables.

The mean scores under the three doses on the auditory tests, as seen in Figure 2, indicate that only two of the tests showed consistent score decrease with dose increase: sentences and unrelated words. Under 50 micrograms of LSD-25 paired associate scores showed an increase followed by a decrease under the next higher dose. The digit span increased very slightly under 50 micrograms and remained at that level under 100 micrograms.

The probability that the score differences between doses are due to chance fluctuations is given in Table 1. Only some of the observed changes in score were significant. Recall of nonsense syllables showed the greatest percentage

change and was among those tests whose differences were most significant. Digit span, on the other hand, was affected little, if at all, by LSD-25, with a very slight increase in scores on successive testings. The number of correct

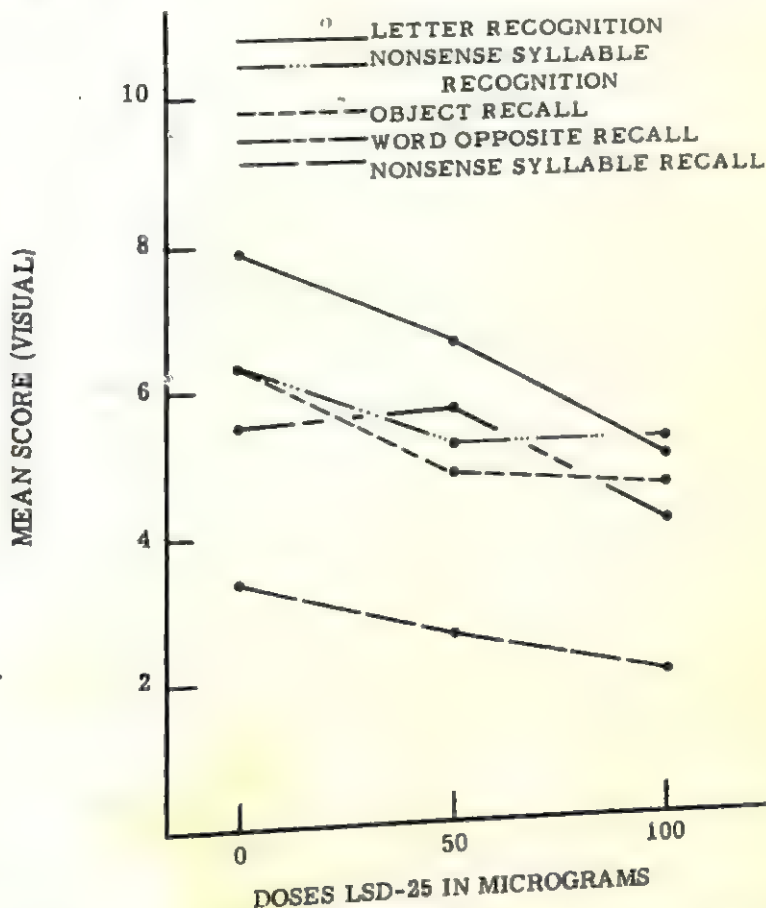


FIGURE 1

Mean number of correct memories on 5 visual recall and recognition tests. Twelve subjects were tested under zero and two doses of LSD-25.

items in each test tended to decrease with increasing dosage, and the number of errors made tended to increase with the dosage increase. Letter recognition, nonsense syllable recall and recognition, paired associate recall, and sentence recall best illustrated this phenomenon. The only increase in errors which was significant at the .05 level or better was that between zero and 100 micrograms on the letter recognition test.

In Figures 3 and 4 the effects of repeated experience with the test are



summarized. Eight subjects were studied. The scores on the first placebo and on the second placebo testing are called pre- and post-LSD-25 placebos, respectively. All the visual and auditory tests showed improvement on suc-

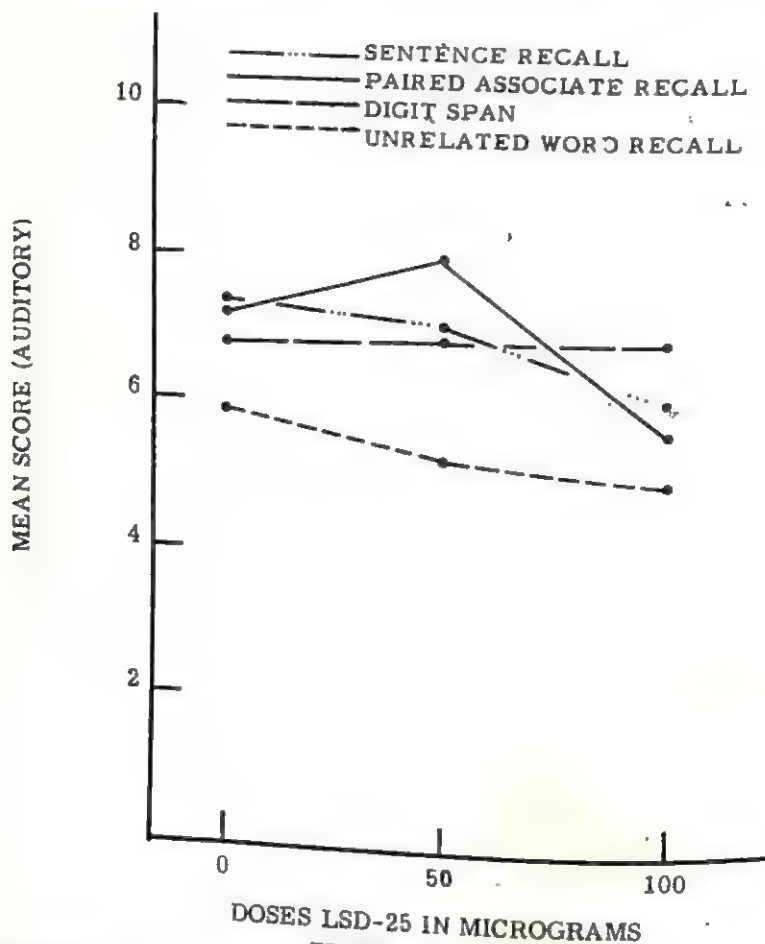


FIGURE 2

Mean number of correct memories on 5 auditory recall and recognition tests. Twelve subjects were tested under zero and two doses of LSD-25.

cessive testings. The second time that subjects received a placebo, with one exception, the scores were the highest of the four experimental days. (Paired associate scores were highest under a 100 microgram dose.) The sentence recall improvement was significant at the .02 level. The changes in performance on the other tests could have occurred by chance more than five times in 100 trials with an  $N$  of eight.

The Spearman rank-order correlations, comparing for each test the sub-

ject's rank position within the group under different doses, are given in Table 2. No significant correlations were found for letter recognition, object recall, and word opposite recall. Nonsense syllable recognition, nonsense syl-

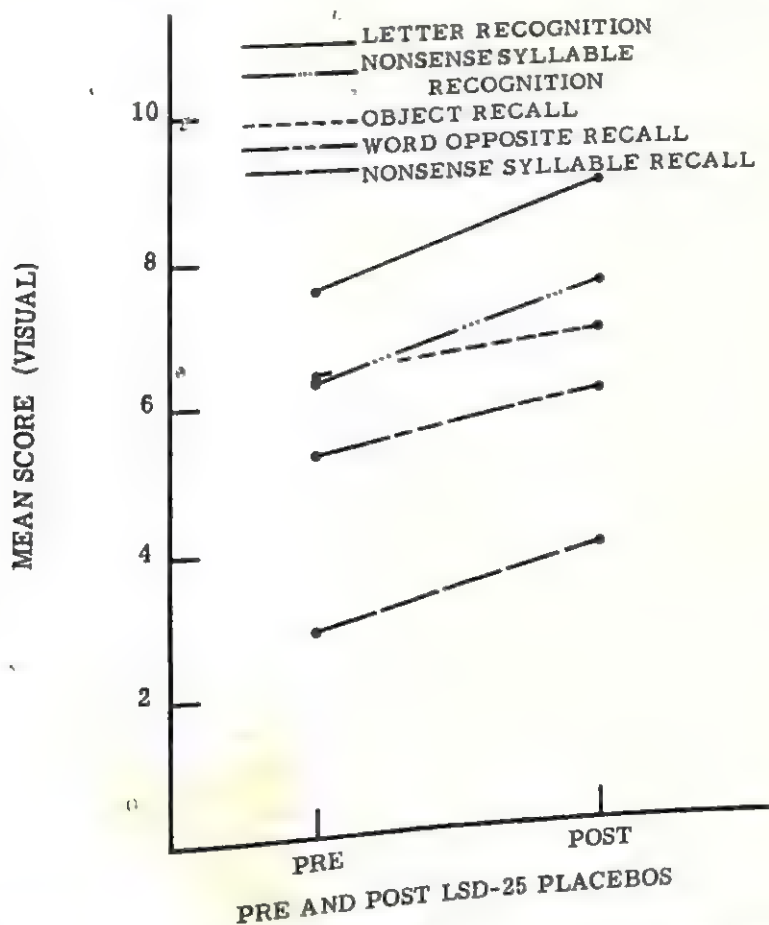


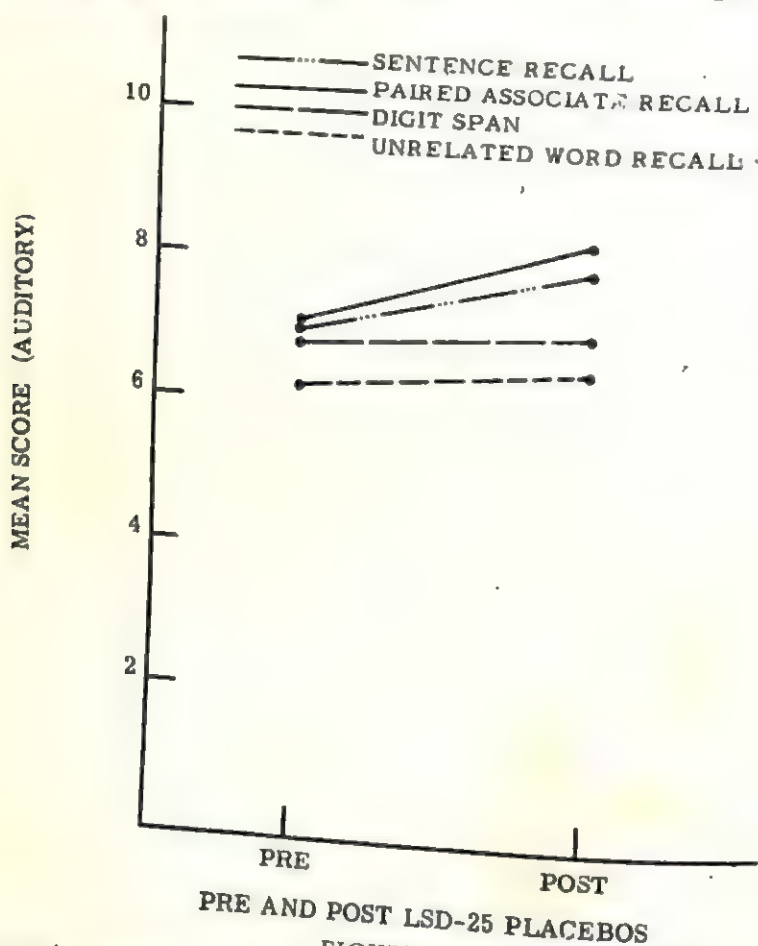
FIGURE 3

Mean number of correct memories on 5 visual recall and recognition tests. Eight subjects were tested under zero dose LSD-25 before and after drug tests.

lable recall, and unrelated word recall showed significant correlations of at least .74 between the two LSD-25 doses. Sentence recall ranks correlated significantly between zero and 100 micrograms and between 50 and 100 micrograms; digit span between zero and 50, and between 50 and 100 micrograms. A significant correlation was obtained for the recall of paired associates between zero and 100 micrograms of the drug.

## D. DISCUSSION

The results show that LSD-25 has a deleterious effect upon the learning of various kinds of visual and auditory stimuli. The situation is reminiscent of the findings of Elliot (2) on latent learning where learning or practice



PRE AND POST LSD-25 PLACEBOS

FIGURE 4

Mean number of correct memories on 5 auditory recall and recognition tests. Eight subjects were tested under zero dose LSD-25 before and after drug tests.

apparently occurred even during poor performance. Although only some of the score decrements were significant and only one pre-post improvement was significant, the trend indicates that steady improvement without the drug was to be expected.

The consistent increase of scores on the post-LSD-25 zero dose over the pre-LSD-25 scores might indicate that there is a practice effect which would

be more significant if a larger population were used. This in turn suggests that the decrements in score under the increasing drug doses are more significant than they appear to be. Where improvement was anticipated on a

TABLE 1  
PROBABILITY OF CHANCE OCCURRENCE OF SCORE DIFFERENCES BETWEEN DRUG DOSES  
(N = 12)

Test	LSD-25 Dosages compared					
	Pre-Drug-Zero vs. 50 Micrograms		Pre-Drug-Zero vs. 100 Micrograms		50 Micrograms vs. 100 Micrograms	
	Dir. of change	P	Dir. of change	P	Dir. of change	P
<i>Visual</i>						
letter recognition*	L	—	L	<.02	L	—
nonsense syllable recognition*	L	—	L	—	N	—
object recall	L	.05	L	<.05	L	—
word opposite recall	H	—	L	.05	L	.05
nonsense syllable recall	L	—	L	.02	L	.05
<i>Auditory</i>						
sentence recall	L	—	L	<.01	L	—
paired associate recall	H	—	L	<.05	L	.05
digit span	H	—	H	—	N	—
unrelated word recall**	L	—	L	—	L	—

*Notes:*

P indicates probability of chance occurrence of difference.

— indicates that P is greater than .05.

L signifies lower score under higher dose.

H signifies higher score under higher dose.

N signifies no change in score.

\* Score corrected for guessing.

\*\* N = 8.

second and/or third administration of the same tests, impairment frequently appeared. If the form of the learning curve were known, drug scores could be compared with anticipated scores after an equivalent number of repetitions. No norms were available, however, and for practical reasons could not be established in this study.

It was expected that, analogizing from Jost's law (11), tests involving familiar items, numbers, words, and letters would most likely be less affected by LSD-25 than tests using unfamiliar items. Although digit span bears this out, letter recognition, word opposite recall, and sentence recall showed impairment under the drug, while an unfamiliar item, nonsense syllable recognition, showed no significant change under the drug. Recall of nonsense syllables, however, showed very significant impairment under the drug. It is not only the material which must be remembered but the method by which



the memory is tested which must be considered. In another of the studies in this series it was found that on the Wechsler-Bellevue Intelligence Scale the subjects under the drug had significantly lowered "information" scores while their digit span score remained constant.

TABLE 2  
SPEARMAN RANK-ORDER CORRELATION OF SUBJECT'S SCORES  
(Three LSD-25 doses are correlated, two at a time;  $N = 12$ )

Test	LSD-25 Dosages correlated					
	Pre-Drug-Zero vs. 50 Micrograms		Pre-Drug-Zero vs. 100 Micrograms		50 Micrograms vs. 100 Micrograms	
	$\rho$	$P$	$\rho$	$P$	$\rho$	$P$
<i>Visual</i>						
letter recognition	.14	—	.11	—	.43	—
nonsense syllable recognition	.19	—	.27	—	.77	<.01
object recall	.20	—	.51	—	.06	—
word opposite recall	-.09	—	.28	—	-.01	—
nonsense syllable recall	.26	—	.49	—	.84	<.01
<i>Auditory</i>						
sentence recall	.29	—	.66	<.05	.66	<.05
paired associate recall	.14	—	.74	<.01	-.08	—
digit span	.65	<.05	.52	—	.69	<.05
unrelated word recall*	-.08	—	-.08	—	.74	<.05

Notes:

$\rho$  is the Spearman rank order correlation coefficient.

$P$  indicates the probability that the correlation could occur by chance.

— indicates that  $P$  is greater than .05.

\*  $N = 8$ .

If we consider that adults are always learning number sequences, addresses, telephone numbers, and train schedules, the fact that there was no significant change of digit span under the drug could be taken to mean that the subjects are quite familiar with this learning situation and have a practiced skill in this type of learning that is very difficult to disrupt. With the other "familiar" items the subjects have not usually had such consistent, every-day practice. However, an *a priori*, non-empirical decision as to which items are more difficult, complex, or unfamiliar, may not be feasible.

The findings reported in Table 2 suggest that the scores on tests of memory under one dose tend to be independent of the scores made under another dose. The significant correlations between 50 and 100 microgram score ranks on the nonsense syllable recognition test, and between the score ranks under zero and 50 micrograms, and those under 50 and 100 micrograms on the digit span test indicate reliability of these tests since the average scores were the same or only slightly different for these comparisons.

The significant correlations between the two drug doses on nonsense syllable recall, and between zero and 100 micrograms on sentence recall and paired associate recall do show that on these tests prediction of a subject's position within the group is possible with an acceptable degree of success. Sentence recall and unrelated word recall also showed significant correlations between the two drug doses. However, the score changes were not significant, and the correlations are of little value for predicting. They do indicate that the tests are reliable.

The results show that LSD-25 has a deleterious effect upon the recall and recognition of various kinds of visual and auditory stimuli. They do not indicate whether the impairment is due to the inability to learn, the inability to retain, or the inability to express what has been remembered. Perhaps the impairment results from inadequate attention to the stimuli. Jarvik *et al.* (5) demonstrated that LSD-25 does reduce the ability of a subject to attend and to concentrate.

#### E. SUMMARY AND CONCLUSIONS

Twelve non-psychotic adults were tested with nine tests of recognition and recall under a 50 and 100 microgram dose of lysergic acid diethylamide, and under a zero placebo before and after the LSD-25 tests. The following tentative conclusions can be drawn:

1. One hundred micrograms of LSD-25 significantly impaired recognition or recall on six of nine tests, when compared to placebo performance.
2. Fifty micrograms of LSD-25 produced an insignificant effect on recall and recognition.
3. Only sentence recall showed significant improvement from pre- to post-LSD-25 placebo sessions. All tests showed some improvement, however.
4. It is possible to predict a subject's rank position under 100 micrograms from his position under zero micrograms on sentence recall and paired associates recall tests. It is also possible to predict his position under 50 micrograms on nonsense syllable recall. Scores will be different, however, under the two doses.
5. Digit span and nonsense syllable recognition are reliable tests showing no significant changes under the drug.

#### REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psych.*, 1955, **39**, 3-60.
2. ELLIOT, M. H. The effect of change "drive" on maze performance. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 185-188.

3. GARRETT, H. E. *Statistics in Psychology and Education*. New York: Longmans, Green, 1953. Pp. 451.
4. HILGARD, E. R. Methods and procedures in the study of learning. In Stevens, S. S. (Ed.) *Handbook of Experimental Psychology*. New York: Wiley, 1951. (Pp. 517-567.)
5. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25): IV. Effect on attention and concentration. *J. of Psychol.*, 1955, **39**, 373-383.
6. SAVAGE, C. Lysergic acid diethylamide (LSD-25): A clinical-psychological study. *Am. J. Psychiat.*, 1952, **108**, 896-900.
7. SLOANE, B., & DOUST, J. W. L. Psychophysiological investigations in experimental psychoses: Results of the exhibition of d-lysergic acid diethylamide to psychiatric patients. *J. Ment. Sci.*, 1954, **100**, 129-144.
8. THORNDIKE, E. L., & LORGE, I. *The Teachers Word Book of 30,000 Words*. New York: Teachers College Bureau of Publications, 1944. Pp. 274.
9. WILCOXON, F. *Some Rapid Approximate Statistical Procedures*. New York: American Cyanamid Co., 1949. Pp. 16.
10. WELLS, F. L., & RUESCH, J. *Mental Examiner's Handbook*. New York: Psychological Corp., 1945. Pp. 211.
11. WOODWORTH, R. S. *Experimental Psychology*. New York: Holt, 1938. Pp. 860.

*Neurology Department*

*Mount Sinai Hospital*

*Fifth Avenue and 100th Street*

*New York City 29*

## LYSERGIC ACID DIETHYLAMIDE (LSD-25): VII. EFFECT UPON TWO MEASURES OF MOTOR PERFORMANCE\*

*Departments of Medicine, Neurology, and Psychiatry, The Mount Sinai Hospital,  
New York City*

H. A. ABRAMSON, M. E. JARVIK, AND M. W. HIRSCH<sup>1</sup>

### A. INTRODUCTION

The psychic effects of lysergic acid diethylamide are so dramatic (4, 7, 10, 15) that one tends to overlook possible psychomotor changes. Our prior experiences showed that LSD-25 produces subjective disturbances in coördination and body control (1). Studies with other drugs have shown that both caffeine (5, 8) and tobacco (3, 17) increase tremor as measured with some form of a stylus and hole apparatus. Pursuit meter performance undergoes impairment when subjects are under the influence of alcohol and when they are at altitudes above 13,000 feet (9, 11, 12). Arm-hand coördination is reduced by altitudes above 10,000 feet (13).

The purpose of this paper is to report the effects of lysergic acid diethylamide on performance on two tests of hand-eye coördination: (a) an adaptation of the Dunlap steadiness test, where the subject must remain immobile, and (b) the pursuit rotor test, where the subject must follow a moving target. It was found that LSD-25 interferes, to some extent, with performance on each of these tests.

### B. METHOD

#### 1. Subjects

There were six male and six female volunteer adults who received payment for participation in the experiments. Only those who were considered non-psychotic on a psychiatric interview and a battery of clinical psychological tests (1) were acceptable subjects. Intelligence as measured by the Wechsler-Bellevue Intelligence Scale, ranged from average to superior. Eight subjects were in the latter group. The median age was 27, ranging from 21 to 33. Median weight was 135 pounds, ranging from 103-175 pounds.

\*Received in the Editorial Office on January 10, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was aided by a grant from the Geschickter Fund for Medical Research. We are indebted to Sandoz Pharmaceuticals, Incorporated, for supplies of LSD-25 and other compounds.



There were seven graduate students in psychology or allied fields, two housewives, an auditor, a hair-dresser, and a public relations worker.

## 2. Apparatus

The pursuit rotor consisted of a phonograph turntable, 23.2 cm in diameter, which rotated  $33 \frac{1}{3}$  and 78 revolutions per minute. There was a round copper button, 8 mm in diameter, situated 7.5 cm from the center of the turntable. The subject followed this button with another button 8 mm in diameter, at the end of the movable arm of a hinged stylus 25 cm in length.

The steadiness apparatus was an adaptation of that of Dunlap described by Spaeth and Dunham (14). It consisted of a copper plate with two holes, 8 mm and 4 mm in diameter. The subject held the wooden handle of a steel stylus, 23 cm in length and 2.5 mm in diameter. The time in contact with the button was recorded during five continuous 10-second trials by means of an electric timer. The number of contacts and the time in contact with the sides of the holes were also recorded during five continuous 10-second trials.

## 3. Procedure

*a. Administration of drug.* At 9:30 a.m. the subjects received the drug in 75 cc of water. All subjects were tested three times at least one week apart, under three different drug doses. Eight subjects were tested four times. The first and fourth sessions were zero dose placebos where only 75 cc of water was given to the subjects. At the second and third sessions the subjects received 50 and 100 micrograms of lysergic acid diethylamide, respectively. Since LSD-25 is a colorless, tasteless, and odorless compound the subjects could not detect the dose they received. The first placebo experiment served as a control of the effects produced by variables other than the drug. The second placebo helped evaluate the effects of repeated practice in the same experimental situation with the same tests.

Two experimenters gave the drug (and the tests) to the subjects in groups of 3, 4, or 5. The doses among the group usually varied in a given day. The administration took place in an air-conditioned room where the subjects had been pre-tested and interviewed. One-half hour after the ingestion of the drug the subjects, who had eaten no food since the previous evening, ate a light breakfast. No stimulants such as coffee or tea were permitted and smoking was not allowed during the entire test period.

*b. Administration of tests.* The pursuit-rotor and steadiness tests were

administered individually approximately 2 to 2½ hours after administration of the drug. This was within the time range of maximum drug effect (1). They were preceded by a series of group tests of memory, cancellation, arithmetic, and spatial relations. The pursuit rotor test was given while the subject was standing. The instructions were: "You are to try to keep the point of the stylus in contact with the moving button." The subject then attempted to follow the button rotating at 33 1/3 revolutions per minute. The time (in seconds) in contact with the button was recorded every 10 seconds for 50 seconds. Without warning the subject, the experimenter abruptly increased the turntable speed from 33 1/3 revolutions per minute to 78 revolutions per minute. The time was recorded every 10 seconds for 50 seconds.

In the steadiness test the subject sat behind the copper plate and was given the following instructions: "Keeping your arm extended, you are to insert the stylus through the large hole to a depth of about one inch and hold it there without touching the sides of the hole." The number of contacts and the time (in seconds) in contact with the sides of the hole were recorded every 10 seconds for 50 seconds. The subject then held the stylus in the small hole and another five 10-second recordings were taken. The 33 1/3 rpm sub-test on the pursuit rotor and the small hole steadiness tests were added to the battery after the first experimental day, thereby reducing the number of subjects whose data were available for statistical analysis.

*c. Analysis of data.* The average time in contact during the 50-second period under each dose was computed for the pursuit rotor and steadiness tests. The average number of contacts on the steadiness tests was also computed for each dose. The averages for a given test did not include the data of subjects who were not tested at all three doses.

The scores obtained under the first three doses on each of these tests were compared, two at a time, by means of the Wilcoxon non-parametric test of paired replicates (16) to determine whether there were significant differences between any two doses. This technique was also used to determine whether there was any significant improvement occurring with practice on the test. The pre-drug zero scores and the post-drug zero scores were compared.

The Spearman rank-order correlation (6) method related the rank position of the subject's scores under two doses. The three doses were related, two at a time. The correlation indicates the predictability of a rank position under one dose from the rank position under another.

## C. RESULTS

Figure 1 shows the mean number of seconds out of 50 seconds in which the subjects were in contact with the copper button on the pursuit rotor test. The longer the time in contact the better the performance. At 33 1/3 rpm

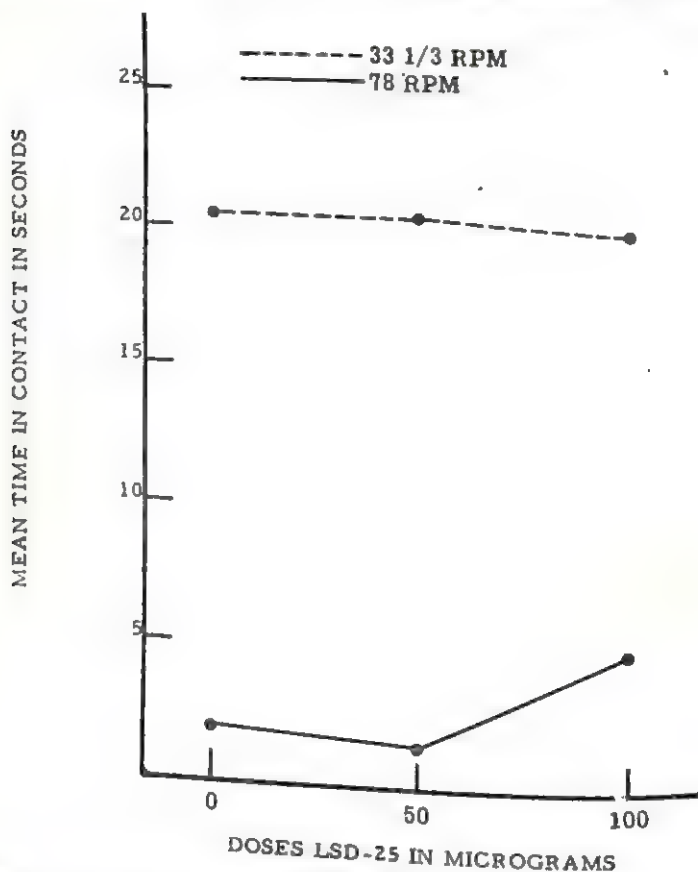


FIGURE 1

Mean time in contact, in 50-second period, with rotating button. Nine subjects were tested under zero and two LSD-25 doses at 33 1/3 rpm and 11 subjects at 78 rpm.

the mean time in contact remained approximately the same under the pre-drug zero and under 50 micrograms of LSD-25, but dropped slightly under 100 micrograms of the drug. At 78 rpm the mean time in contact dropped slightly with doses of 50 micrograms; it increased above the zero dose average with doses of 100 micrograms. All of these changes could have occurred by chance more than 5 times in 100 trials, with the number of subjects used. There were nine subjects tested at 33 1/3 rpm and 11 subjects at 78 rpm under all three doses.

On the steadiness tests the time in contact with the sides of the large and small holes is plotted in Figure 2. The time in contact increased with increasing dosage. The time in contact with the sides of the small hole was always greater than the time in contact with the sides of the large hole,

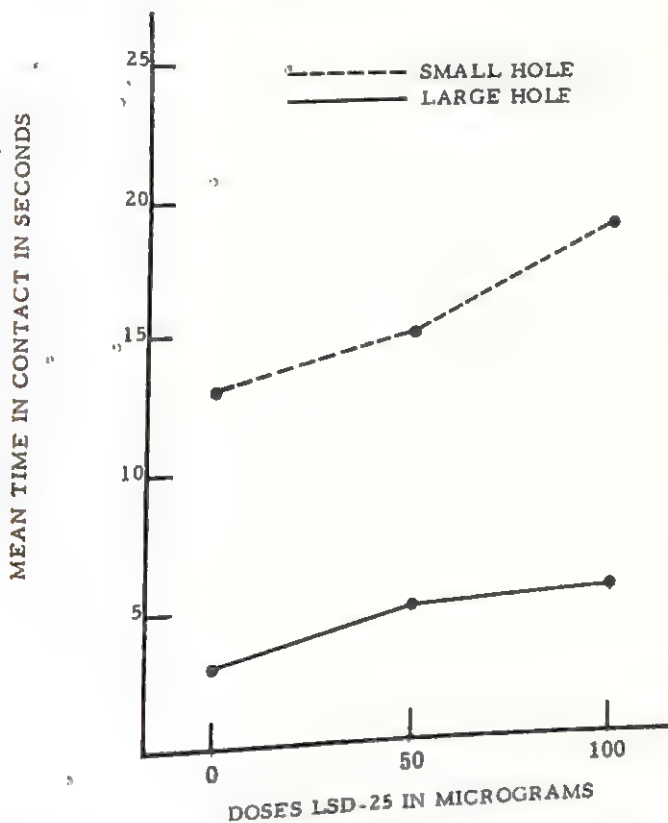


FIGURE 2

Mean time in contact, in 50-second period, with the sides of a large and small hole. Twelve subjects were tested under zero and two LSD-25 doses with the large hole and eight with the small hole.

regardless of dose. None of these changes was significant at the .05 level or better, with the *N* used. Twelve subjects were tested with the large hole and eight with the small hole.

The mean number of contacts made with the sides of the large and small holes appears in Figure 3. There was an increasing number of contacts made with the sides of the large hole with increasing doses. The maximum average number of contacts with the sides of the small hole was made under 50 micrograms of the drug. These changes could have occurred by chance more



than five times in a hundred trials, for the 12 subjects tested with the large hole and the eight subjects tested with the small hole.

Table 1 gives the average scores under the pre- and post-LSD-25 zero dose sessions on the pursuit rotor and steadiness tests and the probability

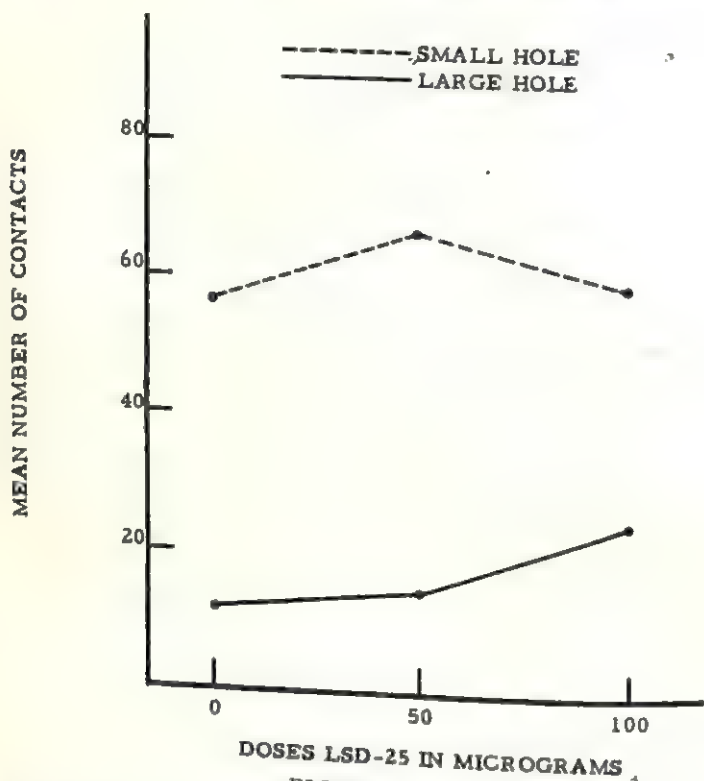


FIGURE 3

Mean number of contacts, in 50-second period, with the sides of a large and small hole. Twelve subjects were tested under zero and two LSD-25 doses with the large hole and eight with the small hole.

that the differences between the two sessions, on each test, could have occurred by chance. The mean time in contact with the copper button increased under the second zero dose experiment at both speeds of the turntable. The mean time in contact and the mean number of contacts made with the sides of the large and small holes both decreased under the second zero dose. The changes which took place on the 33 1/3 rpm and the large hole steadiness tests could have occurred by chance more than 5 times in a 100; the increased time in contact at 78 rpm was significant at better than the .02 level. The decreased time in contact and decreased number of contacts with

the sides of the small hole could have occurred by chance only one time in 100 trials and was therefore a significant decrease.

The relationship between a subject's rank position on a given test under two different doses of the drug is indicated by the size of the Spearman rank-order correlation coefficient. The coefficients relating the scores on each test under the three doses related two at a time appear in Table 2, together

TABLE 1  
COMPARISON OF PRE- AND POST-LSD-25 PLACEBOS  
(Means and Probability of Chance Occurrence of Differences)

Test	Mean score in 50 sec. period		P
	Pre LSD-25	Post LSD-25	
<i>Pursuit Rotor</i>			
Time in contact, in sec.			>.05
a. 33 1/3 rpm*	22.2	28.3	
b. 78 rpm**	2.7	5.4	<.02
<i>Steadiness</i>			
Time in contact, in sec.			>.05
a. large hole**	2.7	0.7	
b. small hole*	14.7	7.8	.01
Number of contacts			>.05
a. large hole**	12.5	4.5	
b. small hole*	58.8	35.7	.01

P indicates probability of chance occurrence of differences between two controls.

\* N = 6.

\*\* N = 8.

TABLE 2  
SPEARMAN RANK-ORDER CORRELATION COEFFICIENTS RELATING SCORES UNDER  
SEVERAL LSD-25 DOSES  
(Three Doses Related, Two at a Time)

Test	Microgram doses LSD-25 compared								
	Pre-drug zero vs. 50			Pre-drug zero vs. 100			50 vs. 100		
	<i>p</i>	<i>P</i>	<i>N</i>	<i>p</i>	<i>P</i>	<i>N</i>	<i>p</i>	<i>P</i>	<i>N</i>
<i>Pursuit Rotor</i>									
1. Time in contact, in sec.									
a. 33 1/3 rpm	.65	—	9	.72	.05	9	.95	.01	10
b. 78 rpm	.63	.05	12	.60	.05	11	.80	.01	11
<i>Steadiness</i>									
1. Time in contact, in sec.									
a. large hole	.40	—	12	.48	—	12	.69	.05	12
b. small hole	.60	—	8	.58	—	8	.83	.01	11
2. Number of Contacts									
a. large hole	.62	.05	12	.70	.05	12	.72	.01	12
b. small hole	.01	—	8	.48	—	8	.36	—	11

Notes:

*p* is the Spearman Rank-Order Correlation Coefficient.

*P* is the probability of chance occurrence of the correlation coefficients.

*N* is the number of subjects.

— indicates that *P* is greater than .05.

with the probability that the correlations would arise by chance. The coefficients ranged from .01 to .95. On the pursuit rotor tests all but one correlation was significant at the .05 or the .01 level. The correlation between zero and 50 micrograms on the 33 1/3 rpm test was attributed to chance. The highest and most significant correlations exist between the two drug dose scores.

The time in contact on the steadiness tests under zero and under each of the drug doses was not significantly related. However, there was a significant correlation between scores under doses of 50 and 100 micrograms of the drug for both the large hole and the small hole tests. The number of contacts made with the sides of the large hole was significantly related among all three doses; the most significant relationship occurs between the two drug doses. No significant correlations were obtained among the various doses, on the number of contacts made with the sides of the small hole.

#### D. DISCUSSION

The small sample used in this study tends to belittle the significance of the changes due to LSD-25, but the trends pointing towards a drug effect cannot be overlooked. When the pursuit rotor turntable rotated at 33 1/3 rpm the mean time in contact decreased as the dose increased. Yet, when a second placebo was given the mean time increased. Ammons (2) also demonstrated an improvement with practice on the pursuit rotor test with a turntable rotating at 60 rpm. One would expect that repetition with this test would show progressively higher, not lower, scores. It seems that LSD-25 tends to interfere with expected learning progress in addition to its impairing the original level of performance. The augmented performance under the post-drug zero testing as compared with the pre-drug zero testing suggests that latent learning occurred during the intermediate drug tests.

At 78 rpm the practice effect was a significant one and was apparently so great that under 100 micrograms instead of the expected impairment, improvement resulted. The test is a fairly reliable one since most of the correlations relating the rank position from one dose to another were at least .60 and significant. The changes in score which occurred under LSD-25 seem more significant in view of the correlations. The highest and most significant correlations occurred between the scores under the two drug doses. If there were no change under LSD-25 the correlations between zero scores and a drug score should be about the same as the correlations between two drug dose scores. The significant correlations also mean that the prediction

of a subject's rank from one dose to another is possible, with an acceptable degree of error.

The results on the two tests of steadiness are similar to the pursuit rotor findings. The practice effect which occurred on the post-drug zero test as compared to the pre-drug zero test leads one to expect improvement on trials subsequent to the first trial. Since the opposite phenomenon occurred it can be said that LSD-25 makes people unsteady and also prevents them from improving their steadiness on the two hole steadiness test. The small hole test, which is the more difficult of the two, improved significantly on both the time in contact and the number of contacts. There was also improvement under 100 micrograms in comparison to the scores under 50 micrograms. This can be explained in view of the significant practice effects. The practice which took place apparently was greater than the detrimental effects of LSD-25, and the score was not as low as anticipated.

The large hole test, aside from showing no significant improvement with repeated testings, was more reliable than the small hole test in terms of the more significant rank-order correlations between different doses on the total number of contacts made. The difference in correlation between the time in contact with the large hole and with the small hole was not as marked. Only the correlation between the two drug doses was significant.

#### E. SUMMARY AND CONCLUSIONS

Twelve adults took two tests of hand-eye coordination, a pursuit-rotor test and a modified Dunlap steadiness test, under 50 and 100 micrograms of lysergic acid diethylamide and under zero doses both before and after the LSD-25 experiments. The following tentative conclusions are made:

1. With a sample of 12 there were no statistically significant differences between the zero dose scores and the drug scores, or between the two drug scores. The tendency was for increasing impairment with increasing doses.

2. Practice with the tests (no drug) showed that on the more difficult tests (78 rpm on the pursuit rotor, and the small hole on the steadiness test) there was significant improvement in the performance. On the 33 1/3 rpm and the large hole tests there also seemed to be improvement, but it was not significant.

3. LSD-25 tends to interfere with the normally anticipated improvement under subsequent testing with the same tests. It is suggested that some improvement occurs despite the drug. This improvement tends to counteract the effects of the drug on test performance.

4. The more reliable test was the pursuit rotor. The highest correla-



tions existed between the two drug doses, on each sub-test of the pursuit rotor and steadiness tests.

## REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide: I. Physiological and perceptual responses. *J. of Psychol.*, 1955, **39**, 3-60.
2. AMMONS, R. B. Acquisition of motor skill: II. Rotary pursuit performance with continuous practice before and after a single rest. *J. Exp. Psychol.*, 1949, **37**, 393-411.
3. CARVER, D. J. The immediate psychological effects of tobacco smoking. *J. Comp. Psychol.*, 1922, **2**, 279-302.
4. DE SHON, H. J., RINKEL, M., & SOLOMON, H. C. Mental changes experimentally produced by LSD (d-Lysergic acid diethylamide tartrate). *Psychiat. Quart.*, 1952, **26**, 33-53.
5. FOLTZ, E. E., IVY, A. C., & BARBORKA, C. J. The use of double work periods in the study of fatigue and the influence of caffeine on recovery. *Amer. J. Physiol.*, 1942, **136**, 79-86.
6. GARRETT, H. E. *Statistics in Psychology and Education*. New York: Longmans, Green, 1953. Pp. 453.
7. HOCH, P. H., CATTELL, J. P., & PENNES, H. H. Effects of mescaline and lysergic acid (d-LSD-25). *Am. J. Psychiat.*, 1952, **108**, 579-584.
8. HOLLINGWORTH, H. L. The influence of caffeine on mental and motor efficiency. *Arch. of Psychol.*, 1912, **3**, (No. 22), 1-166.
9. JELLINEK, E. M., & MCFARLAND, R. A. Analysis of psychological experiments on the effect of alcohol. *Quart. J. Stud. Alcohol*, 1940, **1**, 272-371.
10. LEVINE, A., ABRAMSON, H. A., KAUFMAN, M. R., MARKHAM, S., & KORNETSKY, C. Lysergic acid diethylamide (LSD-25): Effect on personality as observed in psychological tests. (In preparation.)
11. MCFARLAND, R. A. The psychological effects of oxygen deprivation on human behavior. *Arch. of Psychol.*, 1932, **22**, (No. 145). Pp. 135.
12. MELTON, A. W. (Ed.) *Apparatus Tests*. Washington, D. C.: U.S. Govt. Printing Office, 1947. Pp. xxi + 1056. (AAF Aviat. Psychol. Prog. Res. Rep., No. 4).
13. RUSSELL, R. W. The effect of mild anoxia on simple psychomotor and mental skills. *J. Exp. Psychol.*, 1948, **38**, 178-187.
14. SPAETH, R. A., & DUNHAM, G. C. The correlation between motor control and rifle shooting. *Amer. J. Physiol.*, 1921, **56**, 249-256.
15. STOLL, W. A. D-lysergic acid diethylamide (LSD-25). A phantasticum of the ergot group. *Schweiz. Archiv. f. Neurolog. u. Psychiat.*, 1947, **60**, 1-45.
16. WILCOXON, F. *Some Rapid Approximate Statistical Procedures*. New York: American Cyanamid Co., 1949. Pp. 16.
17. WINSOR, A. L., & RICHARDS, S. J. The development of tolerance for cigarettes. *J. Exp. Psychol.*, 1935, **18**, 113-120.

133 East 58 Street  
New York City 22

## LYSERGIC ACID DIETHYLAMIDE (LSD-25): VIII. EFFECT ON ARITHMETIC TEST PERFORMANCE\*

*Departments of Medicine, Neurology, and Psychiatry, The Mount Sinai Hospital, New York City*

M. E. JANVIK, H. A. ABRAMSON, M. W. HIRSCH, AND A. T. EWALD<sup>1</sup>

### A. INTRODUCTION

Adult subjects under the influence of lysergic acid diethylamide have exhibited hesitancy, indecision, and impaired abstract thinking according to clinical observations. For example, Rinkel *et al.* (12) emphasized these changes. Others (3, 13, 14) have observed increased distractibility and difficulty in concentrating in lysergized subjects. By means of simple cancellation tests and short memory tests the present investigators have quantitatively demonstrated that subjects under the drug are not able to concentrate and remember as well as in their non-drug state (7, 8). It seemed likely that drugs would affect performance on solutions of simple arithmetic problems, if they would affect any of the functions involved in the solution of these problems. These include a complex of many intellectual variables, including attention, concentration, old memory, and new memory.

Altitudes of 13,000 feet or greater have impaired performance on simple addition and subtraction problems (5, 10, 11). Studies with tobacco show an increase in errors with smoking on addition and multiplication problems with no clear-cut effect on speed (2). Another study (6) reported a loss of speed in the case of non-smokers and a gain for smokers in doing continuous mental addition. Unpublished results obtained in this laboratory (9) on the influence of LSD-25 on the Wechsler-Bellevue Intelligence Scale indicate a significant impairment on the arithmetic subtest which involves solution of verbal arithmetical problems.

It is the purpose of this paper to investigate the influence of lysergic acid diethylamide on the solution of simple combinations of addition and subtraction problems, and to discover whether there is a change in performance from one dose level of LSD-25 to another.

\*Received in the Editorial Office on January 10, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>This investigation was aided by a grant from the Geschickter Fund for Medical Research. We are indebted to Sandoz Pharmaceuticals, Incorporated, for supplies of LSD-25 and other compounds.

micrograms, and under a 100 microgram dose the average scores were lowest. The Form *B* scores were always slightly lower than those of Form *A*, also decreasing with increasing dosage.

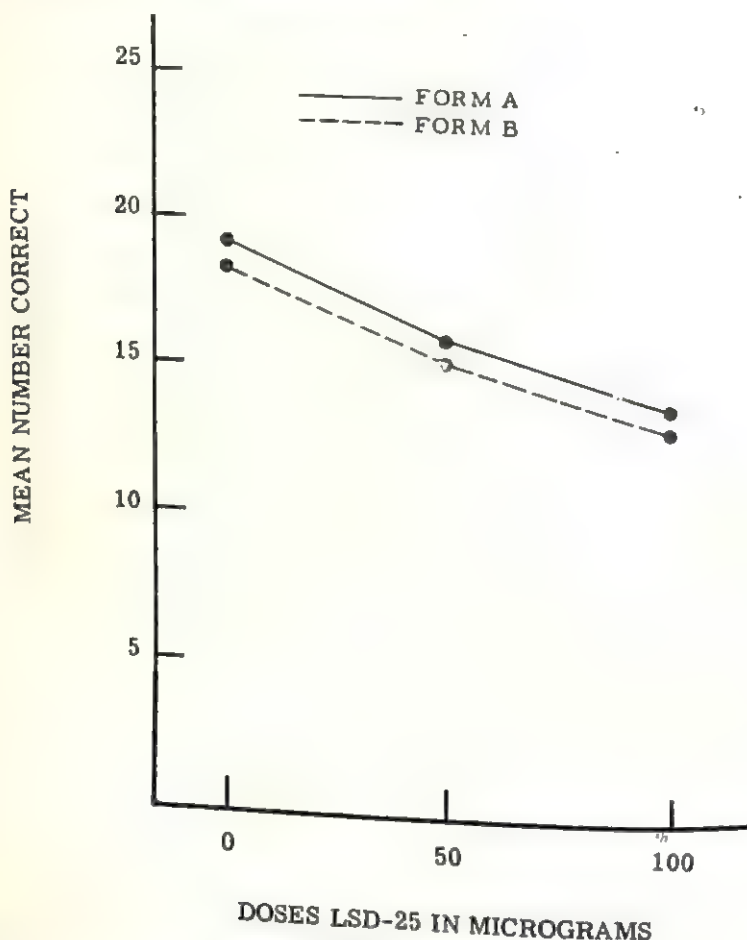


FIGURE 1

Mean number of correct solutions of simple arithmetic problems by 12 subjects tested under zero and two doses of LSD-25

The average number of errors was less than .50 in all three dose groups. There was no significant difference among the groups as tested with the Wilcoxon technique.

Table 1 indicates the probability that the score differences among the three dose levels, taken two at a time, would have occurred by chance. The drop in score from the pre-drug zero to the 50 microgram dose would have occurred by chance less than five times in 100 for Form *A* and Form *B*. Be-

TABLE 1  
SCORE COMPARISONS AMONG LSD-25 AND PLACEBO DOSES  
(Probability of Chance Occurrence of Differences; Relationship Among Scores)  
 $N = 12$

	Compared microgram doses of LSD-25						Post-drug Zero versus B
	Zero versus 50		Zeros versus 100		50 versus 100		
	A	B	A	B	A	B	
Probability of differences occurring by chance*	<.05	<.05	<.01	<.01	>.05	<.05	.05**
( $\rho$ ) Rank order correlation	.69	.67	.51	.44	.58	.65	.02**
(P) Probability of correlation occurring by chance	<.05	<.05	>.05	>.05	<.05	<.05	

\*Performance was impaired under successively higher drug doses. It improved on the second zero as compared to the first.  
\*\* $N = 8$ .



tween zero and 100 micrograms a difference significant at better than the .01 level was obtained for both forms. In comparing the scores under 50 micrograms with those under 100 the difference for Form *A* was not significant but that for Form *B* was significant at better than the .05 level.

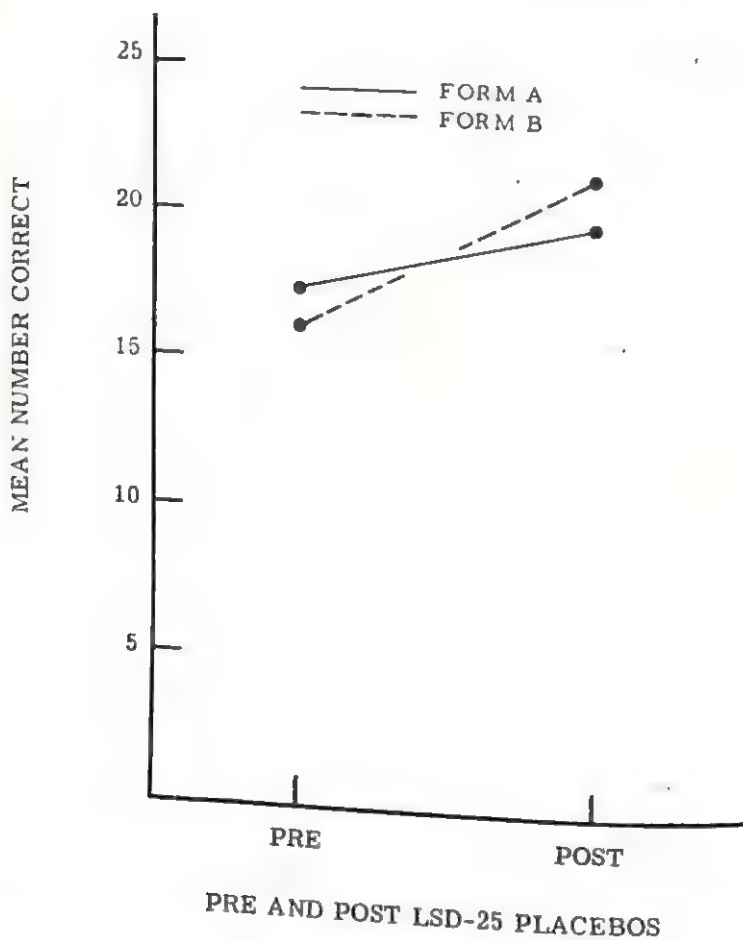


FIGURE 2  
Mean number of correct solutions of simple arithmetic problems by eight subjects tested under zero dose LSD-25 before drug tests and after drug tests

In Figure 2 the results of the pre-drug zero and post-drug zero are compared. The probability of chance occurrence of the differences appears in Table 1. For eight subjects who were available for retesting under a zero dose the mean increased from 17.4 on the first testing with Form *A* to 19.8 on the second testing. The increase is significant at the .05 level. Com-

parison of results on Form *B* demonstrates an average increase, significant at the .02 level, from 16.1 to 20.4 correct solutions.

Table 1 also gives the rank-order correlations among the scores under the first three doses, compared two at a time. For an *N* of 12, correlations between the pre-drug zero and 50 microgram scores and between the 50 and 100 microgram scores were significant at better than the .05 level, for Forms *A* and *B*. These coefficients ranged from .58 to .69. The correlations between pre-drug zero and 100 micrograms were lower and could have occurred by chance more than five times in 100, with the *N* used.

#### D. DISCUSSION

The comparisons among the first three doses become more significant when interpreted in relation to the pre- and post-drug zero sessions. The fact that subjects showed a significant improvement between pre-drug zero and post-drug zero testings indicates that factors which might have operated the first time (such as unfamiliarity with the test, anxiety about the new drug situation, etc.) were present to a lesser extent when the subjects took the last tests. This suggests that if no drug had been given, the two intermediary testings would very likely have shown a progressively increasing number of correct responses. If we knew the effect of practice on this test and could factor it out, then the decrement which appears with increasing drug doses would be even greater and more significant than at present.

The decrease in score on Form *B* in relation to Form *A* on the first three testing sessions was slight when compared to the changes among the various doses. It is interesting, however, that on the second zero dose testing Form *B* scores were higher.

A high rank-order correlation indicates that a subject's position relative to others in the group is similar in two categories which are compared. This type of statistic is unaffected by changes in either the average score or the variability of the group. What it emphasizes is the relationship of members of the group to one another. The data of this experiment show that the highest correlation existed between pre-drug zero and 50 microgram scores. This indicates that LSD-25 tends to disrupt the correlation. The widest discrepancy in rank-order under two doses occurred between zero and 100 micrograms, the two groups which are most different from each other in dose. Thus, higher doses of LSD-25 disrupt the correlation even more than the lower doses.

In order to assess the exact effect of LSD-25 it is necessary to evaluate the mechanisms involved in performance of these simple arithmetic problems.

The subject must not only manipulate numbers in performing the first operation, but he must remember this result for use in executing the second operation. In addition, he must maintain the early formed concept of numbers.<sup>3</sup> The solution of simple arithmetic problems probably demands ability to do abstract thinking, to concentrate and shut out distractions, and to recall numbers. The arithmetic test used in this study appears to be one of the more sensitive tests used by us for detecting the presence of LSD-25, since it was most consistently and most significantly impaired. As a single test it further supports the findings that lysergic acid diethylamide impairs abstract thinking (12), the ability to concentrate (7), and the ability to recall (8) in "normal" adults.

### E. SUMMARY AND CONCLUSIONS

On the basis of 12 non-psychotic adults who were tested under 50 microgram doses of lysergic acid diethylamide, under 100 microgram doses, and under zero doses both before and after the LSD-25 tests with two forms of simple arithmetic problems, the following conclusions can be made:

1. Scores on the tests varied inversely with the drug dose, i.e., the higher the dose, the lower the score.

2. Scores under the various doses were significantly different from each other in five of the six comparisons made. Only the difference between 50 and 100 micrograms on Form A could have occurred by chance more than five times in 100 trials.

3. There was no significant difference in the number of errors made under the several doses of LSD-25.

4. There was a significant practice effect between the first time subjects were tested under a zero dose and the second time they were tested under a zero dose. The latter testing followed the 50 microgram and 100 microgram tests.

5. There was a significant rank-order correlation between zero and 50, and between 50 and 100 micrograms LSD-25 test scores but no significant correlation between zero and 100 micrograms LSD-25 test scores.

### REFERENCES

1. ABRAMSON, H. A., JARVIK, M. E., KAUFMAN, M. R., KORNETSKY, C., LEVINE, A., & WAGNER, M. Lysergic acid diethylamide (LSD-25): I. Physiological and perceptual responses. *J. of Psych.*, 1955, **39**, 3-60.
2. CARVER, D. J. The immediate psychological effects of tobacco smoking. *J. Comp. Psychol.*, 1922, **2**, 279-302.

<sup>3</sup>One subject commented that "While under the higher dose of the drug, I found that the numbers in the problems were completely meaningless and as a result, I was unable to carry out the designated operation."

3. CONDRAU, G. Klinische Erfahrungen an Geisteskranken mit Lysergsaure-diethylamid. *Acta Psychiat. et Neurolog.*, 1949, **24**, 9-32.
4. GARRETT, H. E. Statistics in Psychology and Education. New York: Longmans, Green, 1953. Pp. 451.
5. GELHORN, E., & JOSLYN, A. The influence of oxygen want, hyperpnea, and carbon dioxide excess on psychic processes. *J. of Psychol.*, 1937, **3**, 161-168.
6. HULL, C. L. The influence of tobacco smoking on mental and motor efficiency. *Psycho. Monog.*, 1924, **33**, No. 150, 1-161.
7. JARVIK, M. E., ABRAMSON, H. A., & HIRSCH, M. W. Lysergic acid diethylamide (LSD-25); IV. Effect on attention and concentration. *J. of Psychol.*, 1955, **39**, 373-383.
8. ———. Lysergic acid diethylamide (LSD-25): VI. Effect upon recall and recognition of various stimuli. *J. of Psychol.*, 1955, **39**, 443-454.
9. LEVINE, A., ABRAMSON, H. A., KAUFMAN, M. R., & MARKHAM, S. Lysergic acid diethylamide (LSD-25): Effect on intellectual functioning as measured by the Wechsler-Bellevue Intelligence Scale. (In preparation.)
10. LOUCKS, R. P. An evaluation of various psychological performance tests for altitude chamber research. U.S.A.F. Sch. Aviat. Med., Randolph Field, Rep. No. 202 (1), 17, Aug. 1944. Pp. 22.
11. MCFARLANE, R. A. Psycho-physiological studies at high altitude in the Andes: I. The effects of rapid ascent by aeroplane and train. *J. Comp. Psychol.*, 1937, **24**, 147-188.
12. RINKEL, M., DESHON, H. J., HYDE, R. W., & SOLOMON, H. Experimental schizophrenia-like symptoms. *Am. J. Psychiat.*, 1952, **108**, 572-578.
13. SAVAGE, C. Lysergic acid diethylamide (LSD-25): a clinical-psychological study. *Am. J. Psychiat.*, 1952, **108**, 896-900.
14. STOLL, W. A. D-lysergic acid diethylamide (LSD-25), a phantasticum of the ergot group. *Schweiz. Archiv. für. Neurolog. und Psychiat.*, 1947, **60**, 1-45.
15. WILCOXON, F. Some Rapid Approximate Statistical Procedures. New York: American Cyanamid Co., 1949. Pp. 16.

Neurology Department  
 Mount Sinai Hospital  
 Fifth Avenue and 100th Street  
 New York City 29



## AN APPARATUS FOR RESEARCH IN HUMAN SELECTIVE LEARNING\*

*Department of Psychology, Louisiana State University*

CLYDE E. NOBLE AND FRANCIS J. FARESE<sup>1</sup>

### A. INTRODUCTION

Psychologists interested in the phenomena of human multiple-choice behavior in trial-and-error and discrimination learning situations have employed a variety of laboratory apparatus. Included among these have been memory drums, card sorters, tachistoscopes, and other exposure mechanisms together with an assortment of maze, puzzle-box, key-pressing, and lever-positioning devices. Most of the apparatus described in the literature was developed in connection with rather specific, limited-purpose investigations. Existing laboratory equipment thus provides little flexibility with respect to the alternative experimental designs, manipulable variables, and measures of performance desired in more systematic, theoretically-integrated research programs. Heretofore, simplicity of design was obtained at the expense of operational versatility.

The instrument described in this paper incorporates many of the useful features of previous selective learning apparatus while attempting to minimize their limitations.<sup>2</sup> Essentially it is a semiautomatic combination of the maze and the memory drum. The device was conceived as an elaboration of the methods of Hamilton (1), Hunter (4), and of Yerkes (6). By integrating the multiple-choice and spatio-temporal alternation techniques with the principle of automatic stimulus control, the variety of problems which

\*Received in the Editorial Office on January 11, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>The device reported here was designed by the senior author in 1952 at the Human Resources Research Center. It was constructed by A 3/C J. P. Christopher; S/Sgt. R. B. Card and A 1/C J. B. Downs performed the first experimental studies. In 1953 a research program in selective learning was initiated at Louisiana State University, with the technical assistance of the junior author. The work of L.S.U.'s Human Learning Research Laboratory is currently being supported by grants-in-aid from the University's Council on Research.

<sup>2</sup>In his S-R behavior theory Hull (2) has made a useful distinction between problems of *response-selection* (trial-and-error learning) and problems of *stimulus-selection* (discrimination learning). These basic instrumental learning situations are closely related, however, in that both involve all-or-none choices among discrete alternatives; hence the term *selective* behavior. Typically, this type of learning does not require gradations in force, amplitude, latency, or precision of action as do such perceptual-motor skills as tracking, dexterity, and coordination.

may be studied would seem to be limited only by the ingenuity of the investigator. The apparatus can be adapted for either individual or group experimentation, for paced or self-paced presentation, and for a number of verbal, motor, and perceptual-motor task requirements. It will also serve usefully in investigations of reaction time, work or rate of response, control-display relationships, knowledge of results, generalization, transfer of training, concept formation, and problem solving. One may even employ the equipment as a rating device or as an apparatus test in conjunction with personnel classification and selection programs.

An appropriate generic name is desirable. Technically, the present apparatus is a device for measuring changes in performance attributable to learning. In Greek the word meaning "to learn" or "to acquire the habit of" is *manthano* (μανθάνω). Coupling its root with *metron* (μέτρον) in the conventional manner yields *mathometron*, whose English equivalent is *ma-thor-ē-ter*. Since other types of mathometer (e.g., conditioned-reflex mathometers, pursuit mathometers) are common, subsequent publications will refer to the apparatus described here as the Selective Mathometer.

A number of independent variables are capable of being introduced, and automatic recording of trial, time, correct, error, and space-time data (sequential analysis) are provided. Exclusive of commercial projection, timing, and polygraph equipment, the total cost of parts is approximately \$300. The basic circuit is notably free of tubes, condensers, and resistors, thereby obviating many problems of critical component characteristics. Construction is simple and operation is virtually trouble free.

## B. GENERAL DESCRIPTION

An external view of the Selective Mathometer is shown in Figure 1. From left to right, the visible equipment components are: S's response panel with pushbutton keys, detachable covers, and reward lamp; projection screen and automatic slide projector; operation recorder; E's control unit consisting of: stimulus and reward calibration clocks; electronic timers for controlling reward delay, reward duration, inter-stimulus period, stimulus duration, and self-pacing; correct, error and slide counters; pre-set reward switches, reward indicator, serial-terminal reward switch, power switch, projector switch, stimulus and reward calibration switches, pacing switch, self-pacing criterion switch, manual reward switch, and time criterion switch. Additional components not shown in Figure 1 will be described below.

In normal operation, under conditions of paced, individual administration, E and S face each other with the screen intervening. For paced, group ad-

ministration four *Ss* may be tested simultaneously, projecting on a standard opaque screen with the equipment located behind the *Ss*. The recorder, counter, and reward circuits are, of course, separate for each *S*. This ar-



FIGURE 1  
APPARATUS FOR STUDYING HUMAN MULTIPLE-CHOICE BEHAVIOR:  
THE SELECTIVE MATHOMETER

angement usually requires two *Es* but is quite economical of time, particularly under uniform trial conditions. If counter data are adequate for analytic purposes the recorders can be dispensed with, thus reducing both chart-reading time and material cost.

For experimentation in the L.S.U. laboratory with a single *S*, the apparatus consists of five major units: *S*'s panel, *E*'s panel, relay panel, operation recorder, and projector and screen. These components will be discussed in that order.

### 1. *S*'s Response Panel

This is a semicircular plywood board, set at a 45° angle to the table. Around its perimeter are arranged 19 insulated pushbutton keys which actuate Unimax microswitches mounted under the panel. The keys are spaced exactly 10° apart on a 1 ft. radius from a basal fixation plate for *S*'s

bined time line and event marker, being actuated directly by the projector lamp. The remaining 19 pens are connected to the relay panel by a 20-pair cable. The available chart speeds of the recorder are  $\frac{3}{4}$ ,  $1\frac{1}{2}$ , 3, 6, and 12 in. per hr. or per min., and  $\frac{3}{4}$ ,  $1\frac{1}{2}$ , and 3 in. per sec. For easy chart reading, we constructed a simple hand-winding tape puller consisting of metal guides and spools. A low torque, AC, synchronous motor, equipped with a clutch assembly, provides power for rapid rewinding.

### 5. The Projector

This is a 2in. x 2in. automatic slide projector\* (No. 300 Standard Selectroslide), manufactured by the Spindler and Sauppe Co. of Los Angeles, Calif. It provides remote-control continuous operation featuring removable, cylindrical slide magazines. The capacity is 48 numbered slides per cartridge, and the slide-changing mechanism is practically jam proof. When the projector is controlled by Hunter-Brown timers, there are two independently and continuously variable time parameters: stimulus exposure duration and the inter-stimulus period. The former ranges from .05 sec. to 111 sec., the latter from 1.44 sec. to 111 sec.<sup>3</sup> Control of the inter-stimulus period holds equally well for *E*-pacing and *S*-pacing. The rate of stimulus presentation is jointly determined by these two variables. The inter-trial period is manipulated simply by interposing one or more blank slides between the end of one trial and the beginning of the next.

The translucent display screen was constructed of light blue linen sandwiched between two  $\frac{1}{8}$  in. sheets of clear Plexiglas and mounted in a plywood frame. The stimuli are typically projected from a distance of 4 ft. behind the screen producing an  $8\frac{1}{2}$  in. x 12 in. image.

### C. CIRCUIT FUNCTIONS

A simplified schematic diagram of the apparatus is shown in Figure 2. In order to understand its functioning, let us consider a demonstration experiment in compound trial-and-error learning: the paced acquisition of a 4-link reaction chain by serial reinforcement under two conditions of response availability (5).

Four distinctive stimuli and a cue symbol were prepared as 2in x 2in. slides. Due to the fixed capacity (48 items) of the projector magazines, it was necessary to duplicate each slide eight times. The eight trial cycles were then mounted radially in an invariant order with a blank separating the

<sup>3</sup>The lower limit of 1.44 sec. is the average minimum slide changing time for this particular copy of the Model 300 projector.



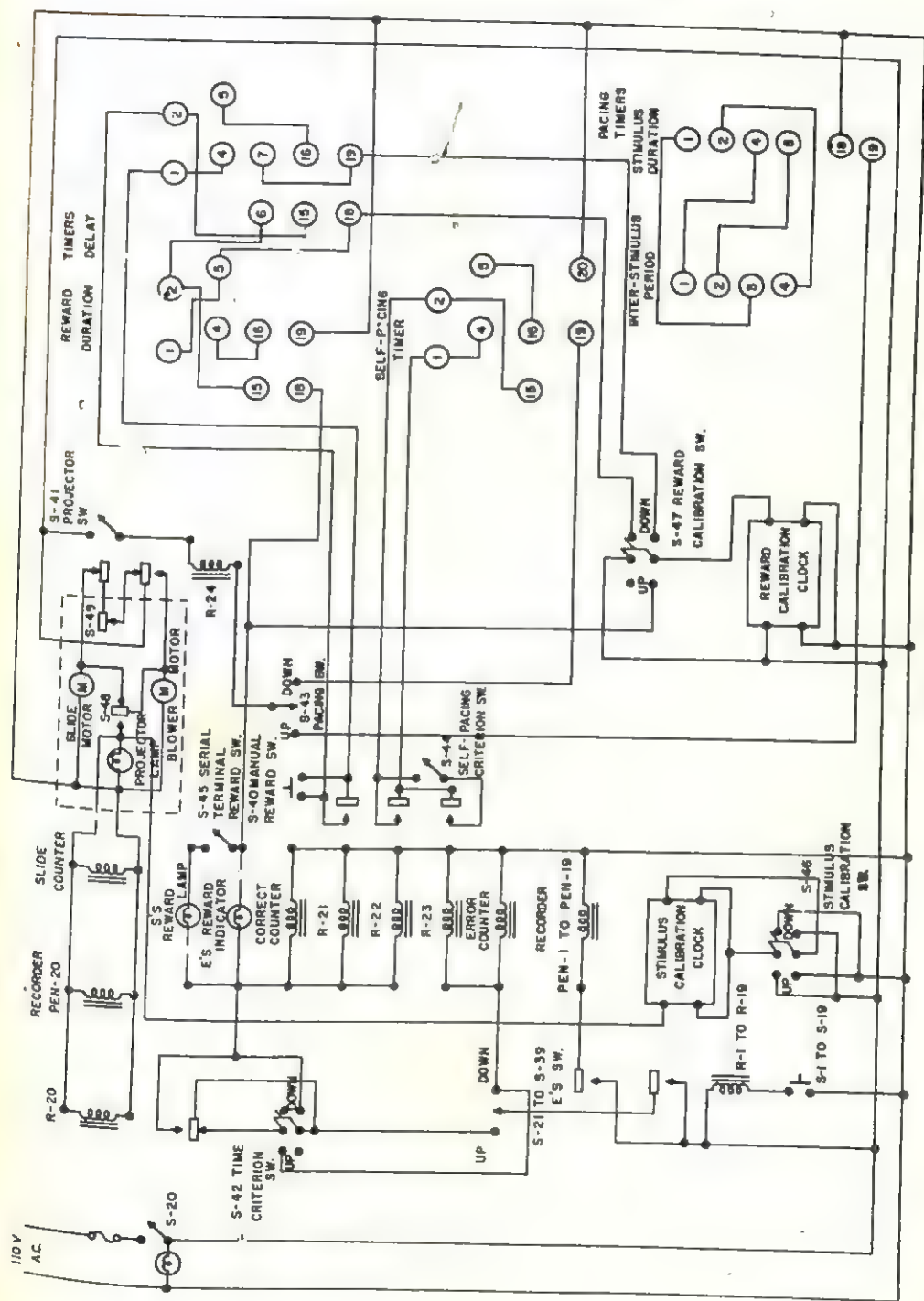


FIGURE 2

SCHEMATIC DIAGRAM OF THE CIRCUIT

For simplicity the internal schematics of the timers and recorder are omitted, together with the Jones plug circuits in S's panel, E's panel, and in the relay panel. The projector circuit is enclosed by dashed lines.

fourth of one and the cue of the next. This constituted a uniform inter-trial interval. The reward timers were set for zero delay and .75 sec. duration. A convenient recorder chart speed for this situation was found to be 6 in. per min. using No. 1720-X charts. The two pacing timers were set for a stimulus duration of 2.00 sec. and an inter-stimulus period of 1.44 sec., giving a presentation rate of one slide every 3.44 sec.

In the present experiment, S's problem was to select from among the 24 permutations of four responses the one sequence which is correct. Learning proceeded by invariant paired associates. The two conditions of response availability were: *four* keys unmasked (4, 8, 12, 16) vs. *nine* keys unmasked (2, 4, 6, 8, 10, 12, 14, 16, 18). In Figure 1 the keys are numbered left to right on S's panel. A condition of nine available responses is shown.

Let the correct sequence for a given subject be:

$$\begin{array}{lcl} S_1 & \longrightarrow & R_8, \\ S_2 & \longrightarrow & R_{16}, \\ S_3 & \longrightarrow & R_{12}, \\ S_4 & \longrightarrow & R_4, \end{array}$$

where the stimulus subscripts are ordinal, the response subscripts nominal. Referring now to Figure 2, we note the following events. First of all, *E* turns on power switch S-20 and the power switches for the reward and pacing timers. *E* next sets the reward delay timer for 0 sec. and the reward duration timer for .75 sec. The stimulus duration and inter-stimulus period timers are set for 2.00 sec. and 1.44 sec. respectively. Time criterion switch S-42 is placed in the up position, allowing only those correct responses which occur while the stimulus is present to be rewarded. Serial-terminal reward switch S-45 is closed for the serial reinforcement procedure. Pacing switch S-43 is placed in the up position for *E*-pacing (standard paced task). Then *E* closes projector switch S-41, initiating the automatic slide sequence.

In order to calibrate the four time variables, *E* performs the following operations. For stimulus duration, Switch S-46 is placed in the up position; for the inter-stimulus period, the down position is used; for reward duration, Switch S-47 is placed in the up position; and for reward delay, the down position is used. The stimulus and reward times are read directly from the two clocks.

When the stimulus duration and inter-stimulus period timers are properly calibrated, Contacts 18 and 19 on the stimulus duration timer automatically close for 2.00 sec. and open for 1.44 sec. in alternation. As these contacts close current flows through S-43 to Relay 24. When Relay 24 is actuated the slide motor within the projector is energized. This causes a slide to

be withdrawn from the cartridge and positioned in front of the projector lamp. At that moment, by means of a cam which controls stimulus onset Microswitch S-48, the projector motor stops as the lamp is lighted. After 2.00 sec. (i.e., when Contacts 18 and 19 open with no current going to Relay 24) current flows through projector motor Microswitch S-49. This energizes the slide motor which, by means of another cam arrangement, opens S-48 and turns the lamp off. At the moment of stimulus offset the slide is replaced in the cartridge.

With the disappearance of the cue slide from the screen, and just before Stimulus  $S_1$  appears,  $E$  places reward Switch S-21 (for  $R_8$ ) in the up position. Note that all other 18 switches S-22 to S-39 inclusive are down. This prepares the circuit from S's Key S-1 through Relay 1, recorder Pen 1 (for  $R_8$ ), S-21, time criterion Switch S-42, the contacts of Relay 20, the correct counter, Relay 21, and the reward timers to S's reward lamp and  $E$ 's indicator lamp. Provided Key S-1 is now selected by the subject before S-48 is again automatically opened (i.e., within 2 sec.), the foregoing loop will be completed thereby counting one correct response, recording  $S_1$  and  $R_8$  in place, time, and duration, rewarding S's response, and indicating visually to  $E$  that S has been rewarded. The delay and duration of reinforcement are controlled by the pre-set reward timers as discussed above.

S-40 is a manual reward switch, available to  $E$  for introducing terminal reinforcement following reaction chains of varying length. It is not used in the present example.

As the illumination of the first stimulus ( $S_1$ ) fades from the screen,  $E$  places S-39 (for  $R_{10}$ ) in the up position, simultaneously returning S-21 to the down position. Assume that the subject responds quickly enough but *incorrectly* to the second stimulus ( $S_2$ ) by selecting Key S-2 (for  $R_{12}$ ). Current will now flow from S-2 through Relay 2, Pen 2 (for  $R_{12}$ ), and through S-22 to the error counter. Since Key S-19 was not selected, Response  $R_{12}$  is properly evaluated as erroneous and no reward is given. This sequence of events is continued trial by trial until either some performance criterion is attained or a constant trial amount or time period is consumed. The manipulation of this representative independent variable has been shown to result in significant differences in performance (5).

#### D. SUMMARY

1. This paper has described a laboratory apparatus of flexible design and simple construction for experimental research in human selective (multiple-choice) learning.

2. The equipment was developed in order to provide reference apparatus for the systematic study of a wide range of patterned, sequential behavior in which the component acts are of a probabilistic or all-or-none character. The primary interest is an experimental analysis of trial-and-error and discrimination learning phenomena.

3. The apparatus has proved adaptable for individual and group experimentation, for paced and self-paced presentation, and for verbal, motor, and perceptual-motor tasks. It may also be employed in investigations of reaction time, work or rate of response, control-display relationships, knowledge of results, generalization, transfer of training, concept formation, and problem solving. One may even use the equipment as a rating device or as an apparatus test in conjunction with personnel classification and selection programs.

4. A number of independent physical variables are capable of being introduced, together with automatic recording of trial, time, correct, error, and space-time data. In extensive use tests operation has been virtually trouble free.

5. Circuit functions were described in connection with a demonstration experiment in human trial-and-error learning. The manipulation of a representative independent variable was illustrated.

6. To describe the present apparatus for technical purposes the term *Selective Mathometer* was coined. In general, a mathometer is any device for measuring learning phenomena.

#### REFERENCES

1. HAMILTON, G. V. A study of trial and error reactions in mammals. *J. Anim. Behav.*, 1911, 1, 33-66.
2. HULL, C. L. *A Behavior System*. New Haven: Yale Univer. Press, 1952.
3. HUNTER, T. A., & BROWN, J. S. A decade-type electronic interval-timer. *Amer. J. Psychol.*, 1949, 62, 570-575.
4. HUNTER, W. S. The temporal maze and kinesthetic sensory processes in the white rat. *Psychobiology*, 1920, 2, 1-17.
5. NOBLE, C. E. Compound trial-and-error learning as a function of response availability ( $N_R$ ). *J. Exp. Psychol.*, 1955, 49, 93-96.
6. YERKES, R. M. A new method of studying the ideational behavior of mentally defective and deranged as compared with normal individuals. *J. Comp. Psychol.*, 1921, 1, 369-394.

Department of Psychology  
Louisiana State University  
Baton Rouge, Louisiana



Published as a separate and in *The Journal of Psychology*, 1955, 39, 485-502.

1923-1953 and 20-60 AGE CHANGES IN MORAL CODES,  
ANXIETIES, AND INTERESTS, AS SHOWN  
BY THE "X-O TESTS"\*<sup>1</sup>

*Department of Psychology, The Ohio State University*

S. L. PRESSEY AND A. W. JONES

It is generally recognized that the American culture has changed substantially in the last 30 years (1), also that attitudes of older and younger people differ. Inter-relationships of these two types of change might be expected. It is also increasingly being appreciated that accumulations of test data over substantial periods of time permit valuable comparisons of various types. The present paper reports latest results from one of the oldest and in certain respects most systematically continued of such studies, yielding age and culture comparisons believed unique.

A. MATERIALS AND CASES

In 1923, O. R. Chambers (3) gave the "X-O Tests" to 779 public school children in Grades 6, 8, 10, and 12 in various communities, mostly in Ohio, and to 955 undergraduate students at Ohio State University. In 1933, 1,080 Ohio State University students were given these same tests and the data analyzed by Walter Buck (2). In 1943, 1,706 students in the same grades in certain Ohio schools and the four undergraduate years at Ohio State University were similarly tested (7). In 1953, 842 students at the same institution were given the same test. In addition, 408 adults ranging in age from the twenties into the sixties were given the first two tests of this "personality inventory."

The "X-O" inventory is a simple and straight-forward little inquiry form which seems to have sundry merits for such wide-range surveys. The first test consists of 125 items beginning "drinking, smoking, flirting, spitting, giggling," and the person taking the test is asked to cross such of these items as he thinks wrong, then circle the one item in the line he thinks worst. Test two begins "loneliness, work, forgetfulness, school, blues" and any items are to be crossed about which the testee has ever worried or felt anx-

\*Received in the Editorial Office on January 18, 1955, and published immediately at Provincetown, Massachusetts. Copyright by The Journal Press.

<sup>1</sup>The work here reported was supported in part by the Ohio State University Development Fund, and by a grant from the U. S. Public Health Service for study of superior older people.

ious; the one item in each line most worried about is then to be circled. In test three the person being tested is to mark those things he likes or is interested in, the first line reading "fortune-telling, boating, beaches, mountains, vaudeville." And the strongest interest in each line is to be circled. In effect each word is a question, as "Do you worry about your work?" The single-word question permits a great number of questions to be covered in a small space of time and blank; in about 20 minutes on three small pages a student answers in this fashion 450 questions. There is a certain colloquial directness about the procedure and a compactness and simplicity about the general idea of the test which seems to make it more readily applicable than more formal and elaborate-looking inventory blanks, for both young children and adults. Though now almost 35 years old, it is believed the test still stands as yielding more scores in a given time and for a given size blank than any other test instrument (5).

In each one of the decade testings it is believed that the students to be tested were selected in about the same fashion—were mostly in courses in psychology. And they were in each instance from the same institution and tested at about the same time of year. But the first survey (in 1923) was in the era of flappers and wonderful nonsense, the second in the midst of the depression, the third in the middle of the Second World War, and the last (in 1953) in the uncertainties of the Korean situation and the cold war. And the adults had been young in these diverse decades: 147 were 60 years old or over, 117 in their fifties, 89 in their forties, and 55 in their thirties. Almost two-thirds were women. Some were club women, some members of a grandmothers club, some wives of the men tested; and the men ranged from a pants-presser to a U. S. Senator! Most were businessmen in several smaller Ohio communities. A good proportion had some college education; most were middle class—a diverse total, got when and where testing was possible, but yielding (as will be seen) interesting results.

## B. RESULTS

Naturally first to be considered is total number of words marked on each test, then per cents marking each item in each group, as explicating the changes in those totals.

### 1. *Decade and Age Changes in Total Words Marked*

Results of the most recent survey, and relations to previous findings, are presented in Figure 1. Most significant appeared to be changes in numbers of somewhat borderland acts—like smoking, flirting, slang,—marked

as wrong. In 1923, eighth grade boys and girls so marked an average of 74 items; the girls in the tenth and twelfth grades marked almost as many; but in college the averages rapidly drop to 53 for senior coeds. A natural

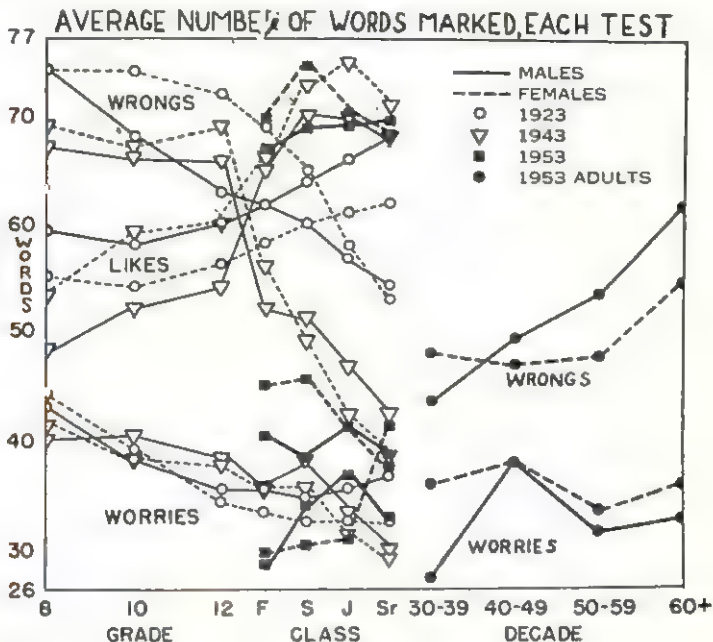


FIGURE 1  
AVERAGE NUMBER OF WORDS MARKED, EACH TEST

interpretation was that as long as the girls remained in the home, its influence remained dominant, but in college, attitudes rather rapidly liberalized. And of course the liberal-minded girls from the more liberal homes may have been the ones who went to college. Meanwhile, the boys, with their presumed greater freedom from home influences, from the eighth grade on considered fewer and fewer of these borderland acts wrong, ending in the college senior year with much the same liberal attitudes as the college senior girls.

By 1943 (as shown by the lines with triangles) home attitudes had apparently liberalized, since eighth grade girls marked five less items as wrong than eighth grade girls 20 years previous, and boys six less. Both sexes show an even greater drop in the college years, the senior coeds to 38. The college senior of '43 appears to have been much more liberal than the senior of '23 ('33 seniors, not shown in this graph, were about half way between). That the two sexes show curves closer together in '43 than '23 suggests

that at the later time they developed more together in moral code. College freshmen in 1953 marked substantially fewer items as wrong than freshmen in 1943. It might be inferred that, if 1953 results had been obtained from the eighth grade on, a substantial liberalizing in home attitudes would have been found. However, the 1953 seniors are only slightly below 1943 seniors. Perhaps change in borderland moral code has slowed down.

The older the adults, the more items they mark as wrong. This looks like increasing conservatism, with age. However, certain cross-readings on the graph seem of much interest. In 1953, men in their 30's marked approximately the same number of these borderland acts wrong as did senior men 10 years previously. Men in their 50's marked about the same number of items wrong as did senior college boys in 1923, when these men were in their twenties. Roughly the adults seem to continue the attitudes reached in their formative period; perhaps they have not become more conservative, but simply not changed!

The curves for number of worries are very different: only a slow drop from eighth grade to college senior year, little change from 1923 to 1953, or from the younger to the older adult years, no sex differences. As will be seen shortly, there are differences on items, but often cancelling each other off. The slight changes in totals may be a happenstance in selection of items for the test such that balancing items, or items with little age change, were included. But another factor is believed important. Worries are not, like moral judgments, so inter-related in codes of behavior that changes in the culture carry them along together and in somewhat consistent ways.

The "likes" curves show yet different patterns: increases from the eighth grade to the college senior year, and especially for the girls in 1943 in high school and college. Here seem shown the increased freedoms and varieties of interests which have opened up for girls in the more recent years. The 1953 data maintain but do not go above the '43 peak. Perhaps the above mentioned trends have slowed down; or perhaps the 35-year-old test does not include newer interests.<sup>2</sup>

Certain further handlings of these data were tried. Thus the 10-90 percentile range was used as a measure of scatter to see whether older adults varied more in total items marked as wrong or worried about. Men 60 and over showed a range of 60; below that age the ranges were in the forties. Women over 60 had a range of 63; below this age, down to the freshman

<sup>2</sup>The increases in interests, and the further finding that every test has some items now marked more than formerly, would seem an answer to any inference that decreases in the more recent testings have simply been a less tendency to mark, perhaps resulting from more test sophistication.



women, ranges varied irregularly from 43 to 59. An increasing or narrowing number of interests in the older years, as one approached retirement, might be hypothecated. But the figures are too irregular to be more positive. The adult males were classed in occupations as clerical and semi-skilled, semi-professional and sales, and professional and administrative. From the first group to the last, there was a slight decline (from 52 to 47) but obviously nothing which could be significant. No appreciable difference between socio-economic classes in number of worries appeared.

#### 8. *Analyses by Item*

To find what was really back of the above-mentioned general trends of the various groups, the per cents marking each item on each test were therefore determined. For this purpose junior and senior results for each sex were averaged to give an average for upper-class students, freshman and sophomore results similarly averaged, also tenth and twelfth grade counts to give an average as of the eleventh grade, and sixth and eighth grade data to give an average as of the seventh grade. Counts on each item for 30- and 40-year-olds were similarly combined, and tallies for adults 50 and over similarly lumped. The totals for men and women were as follows: Fr. So. 201 and 370; Jr. Sr. 132 and 139; ages 30-49, 68 and 76; 50 up, 78 and 186.

The resulting tables showed eight percentages for each of the 125 words of the first two tests and four percentages for each word of the last test, or 2,500 percentages in all (summarizing over 400,000 responses made by the 1,250 cases to the 375 items of the inquiry form). And for comparisons there were 20 percentages for each word, or a total of 7,500 percentages, on the "item norm sheets" from the previous three testings. The most wide-ranging comparisons made possible by the current study were considered to be between the university upper-class group (junior-seniors) in 1923 and 1953 as indicating changing educated young adult attitudes, and between junior-seniors and adults 50 and over in 1953 as contrasting young and old attitudes at that time. The differences between these contrasting percentages were therefore calculated, for each sex. And the total of these differences was used as a rough indication of amount of change in attitude over the 30-year period and roughly 30 years difference in age. Table 1 shows the eight items in the "wrongs" test with largest total difference, plus "war" as the one item showing substantial increases over the 30-year period, and three items illustrative of little change. Also shown are the four worries with largest total differences—except that "school," "lessons," "boys," and

"girls" as worries have been omitted on the ground that they were not really appropriate terms for older adults, and the big drops in per cents of adults marking them not therefore otherwise meaningful. The lower "total differences" on worries than wrongs are to be noted. And again certain other items thought of comparative interest are added.

TABLE 1

Per cents of junior-senior university students in 1923 and 1953, and of adults 50 and over in 1953, marking certain words of the X-O Tests; also, total differences between these successive groups (included are eight wrongs and four worries with greatest total differences plus certain other items mostly with low differences).

	Men			Women			Total diff.
	jun-sen '23 '53	old '53		jun-sen '23 '53	old '53		
<i>Wrongs</i>							
immodesty	70	40	86				
extravagance	61	25	65	80	47	70	135
poolrooms	36	8	33	55	32	63	130
overeating	68	46	63	45	18	61	123
craps	60	16	45	65	30	58	122
slang	43	11	45	61	29	40	116
silliness	68	33	49	48	21	39	111
flirting	47	11	24	61	25	42	104
war	37	65	69	56	19	35	102
thoughtlessness	66	67	62	40	62	63	55
divorce	42	42	44	65	71	67	16
cards	10	5	9	35	33	49	20
<i>Worries</i>							
marriage				5	1	4	16
looks	32	49	31				
popularity	36	44	15	33	56	22	92
business	33	36	26	45	52	18	78
disposition	36	16	50	34	53	12	73
unfairness	26	26	25	14	11	23	69
family	48	47	44	33	37	35	7
friends	50	41	47	55	48	53	16
money	45	33	31	52	61	52	33
	64	60	61	52	52	32	34
				55	65	51	29

How significant are the differences in the above table? And at first glance the words in it seem a rather odd miscellany; what interpretations seem reasonable? It is believed that significance is better indicated by the consistence of all 28 percentages available for each word and coherence with findings for related words, rather than by conventional tests of significance, and that interpretations should be based on such coherence plus relationships with other evidence regarding cultural change. The following pair of charts well illustrates such consistence and coherence. They concern an item—smoking—having a total difference not quite great enough to include it in the table, regarding which a shift from disapproval to ac-

ceptance for women over the past 40 years is well known. Further, the item is also in the likes test. Do disapprovals and likes, for this same item, change consistently? In 1923, over 90 per cent of grade school girls considered "smoking" wrong, but in college there was a rapid drop ending with only some 30 per cent of junior-senior women condemning the indulgence. In 1953, only 5 per cent of upperclass women did so, and the slight drop from freshman-sophomore average suggests that grade school marking (attitudes in '53 homes) would have been more lenient then—as it was somewhat in '43. But the "likes" test shows an increasing number of college women, from '23 to '53, fond of smoking. The men show, relatively, little change in this matter.

Now for three other "minor vices," which are in the table. The table shows substantial decreases in the percentages of college junior-seniors marking "poolrooms" wrong from '23 to '53 (and in '23, 84 per cent of the Grades 6-8 boys and 87 per cent of the girls so marked). The total figures for "craps" are similar. The middle-class American home of 30 years ago evidently considered the smoke-filled poolroom an evil place, and the alley crap-game a lower-class and scorned amusement. Also, slang was not correct or proper. Surely the reader whose recollection goes back near the turn of the century will agree that the conventions of middle-class existence then were indeed circumscribing compared to youthful life now. And the

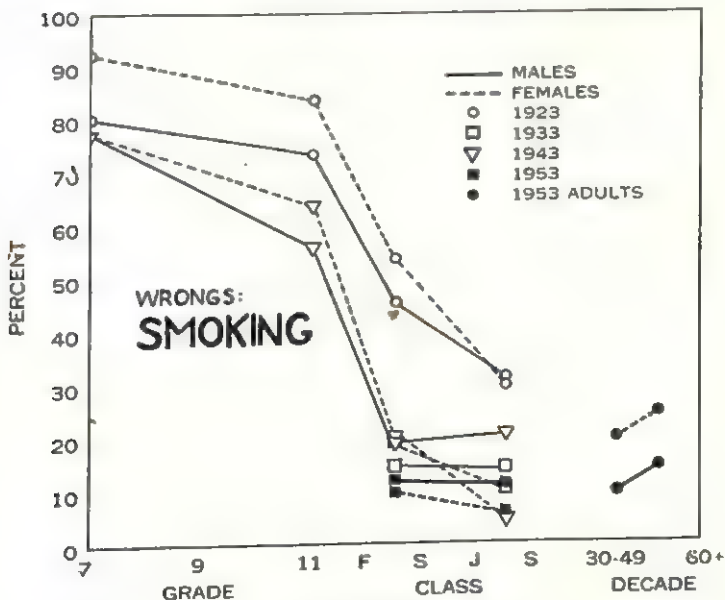


FIGURE 2  
PERCENTAGES IN EACH GROUP MARKING "SMOKING" AS WRONG

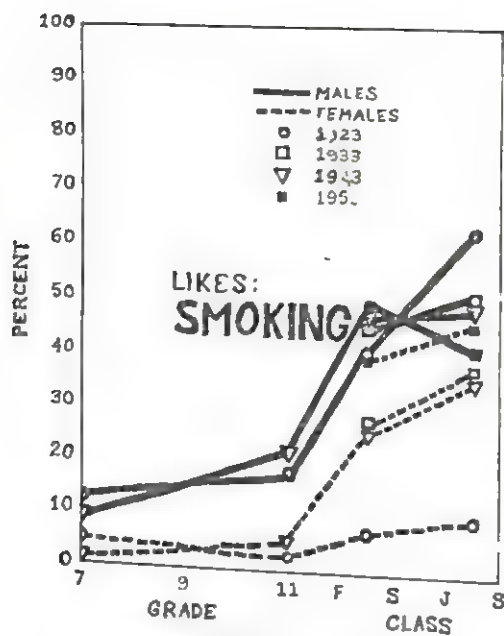


FIGURE 3  
PERCENTAGES IN EACH GROUP MARKING "SMOKING" AS LIKED

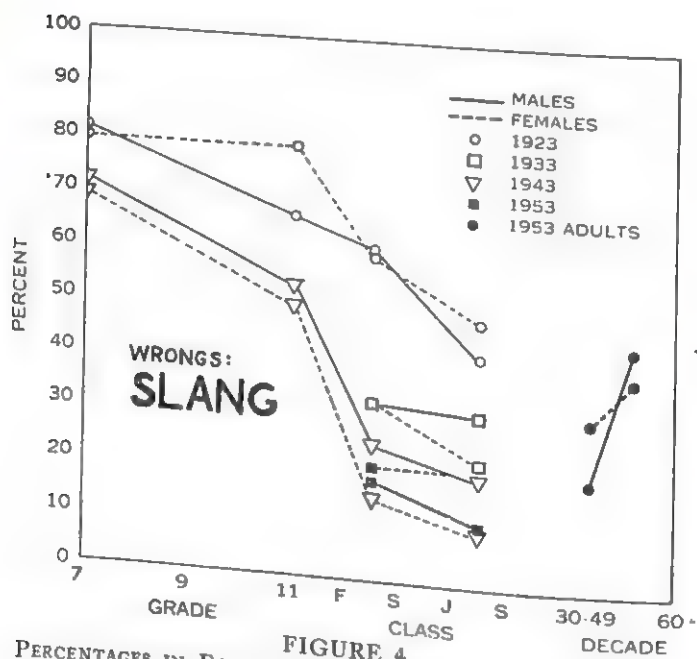


FIGURE 4  
PERCENTAGES IN EACH GROUP MARKING "SLANG" AS WRONG



essential similarity of the graphs for "smoking" and "slang" as wrong—also the graphs, not here included, for "poolrooms" and "craps" and some other like items—suggests that all this is part of a massive cultural trend from strictness to acceptance of certain freedoms of conduct and speech.

Also similar is the graph showing numbers marking "flirting" as wrong. Moreover, "flirting" also appears in the "likes" test, there showing consistent increases. It is amusing further congruence are the increases from '23 to '53 (greatest of any on this test) in percentages frankly marking "kissing" as liked. Further bearing on sex-social attitudes, and in congruence with general knowledge, are the graph lines for "immodesty" as wrong: rises from childhood into adolescence, drops in the college years and from '23 to '53, more condemnation by adults. It seems a fair guess that the less condemnation (shown in the table) of "silliness" now by young people means more acceptance of dalliance. All this is highly consistent with the findings of Kinsey and of others regarding increases in petting and relaxations of sex-social restraints; but the total data would seem to put these last findings in a larger setting. Kinsey (4, pp. 242-5 and 298-302) has stressed influences bearing on sex-social conduct around the time of the first world war—contacts of great numbers of American soldiers then with central European cultures having sexual patterns different from the American, increased knowledge of contraception, popularization of Freud and Havelock Ellis. And he found his greatest changes during and immediately after that war. But the data so far presented in this paper, and much which is to follow, would emphasize more than Kinsey that all these changes are parts of a vast pervasive shift from a restrictive Puritanical to a more rationalistic and naturalistic culture. If the attitudes of Grades 6-8 children in '23 be considered fostered by parents born before 1900, and insofar comparable to Kinsey's group so labelled, then indeed greatest change came "between 1916 and 1930." But the present data show sundry changes occurring thereafter, to a greater extent than Kinsey seems to recognize. If only the X-O Tests had been given in '13, more adequate evidence of change since then would be possible.

The table shows "extravagance" having the second largest total difference, being condemned much less by college students in '53 than '23, and much disapproved by older adults. The chart gives more details: in '23 a majority from grade school through college considered extravagance wrong; in '43 fewer at all these levels so felt, with college upperclass students most indulgent as to spending; and so they were in '53. The figures for "debt" as wrong were quite similar though less striking. Installment buying,

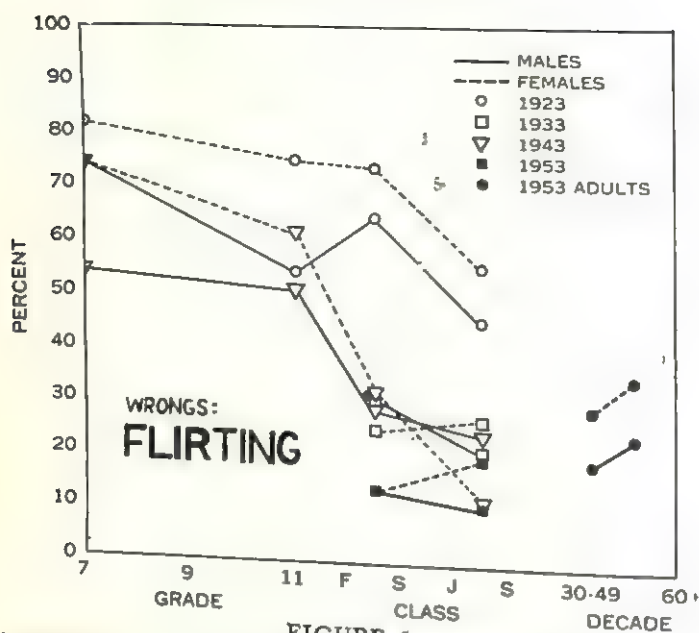


FIGURE 5  
PERCENTAGES IN EACH GROUP MARKING "FLIRTING" AS WRONG

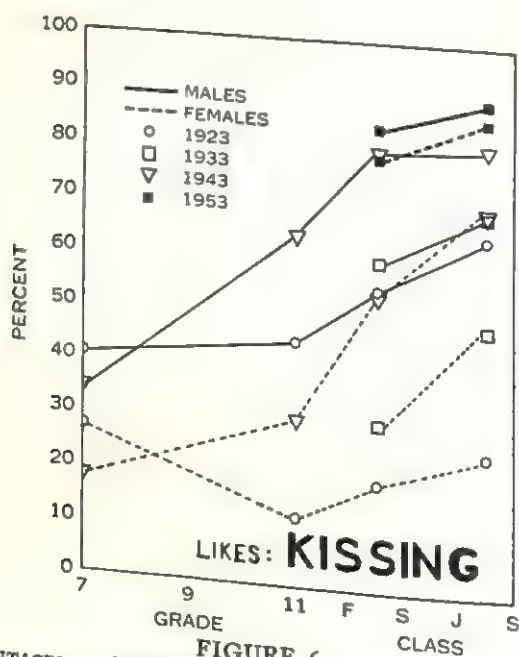


FIGURE 6  
PERCENTAGES IN EACH GROUP MARKING "KISSING" AS LIKED

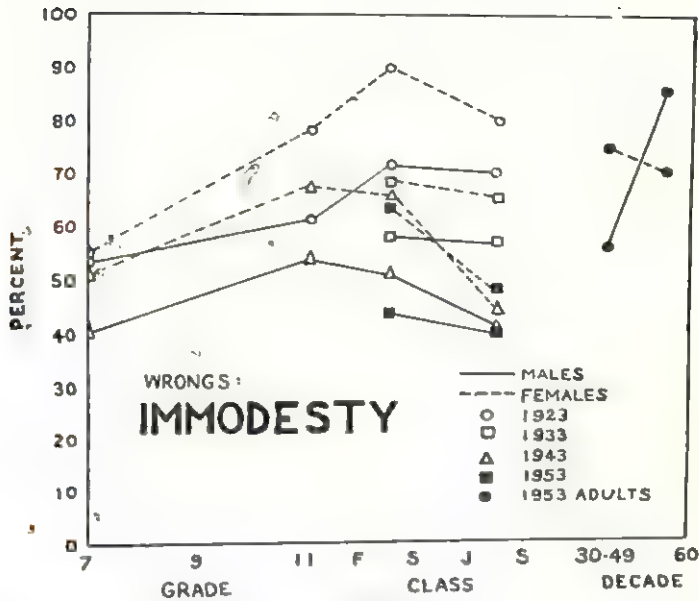


FIGURE 7  
PERCENTAGES IN EACH GROUP MARKING "IMMODESTY" AS WRONG

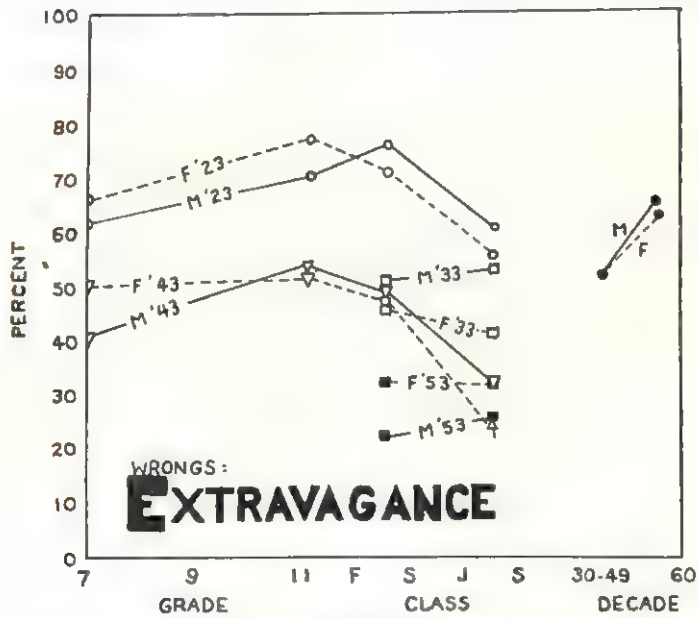


FIGURE 8  
PERCENTAGES IN EACH GROUP MARKING "EXTRAVAGANCE" AS WRONG

inflation, big wages and free spending during and immediately after the war, might account for recent changes—perhaps economic fatalism, during the depression.

Greater condemnation of overeating by older adults than by students was to be expected (as shown in the table) but hardly the less disapproval by students in '53 than '23. However, the second test showed more college girls in '53 than in '23 worrying about overeating. Perhaps dietary attitudes are, now, more health and fashion conscious than moral. Finally, the table shows "thoughtlessness" generally condemned by young and old, in '53 as '23, the total differences being only 16—as compared to 132 for immodesty.<sup>3</sup> There are sundry such items, of elementary considerateness or ethics and not involved with cultural taboos, which do not show change. "Cards" also has a difference total of 16 but for another reason: general acceptance in the groups and for the period covered, though 68 per cent of Grades 6-8 girls in '23 marked this item wrong—doubtless an X-O test in 1883 would have shown general disapproval. The curves for dancing are similar. The low total difference for "divorce" might be interpreted as a stabilization at partial acceptance.

Certain increases in the oldest group are believed indicative of faults it is either most sensitive to or tries especially to guard against. Thus in '53, 34 per cent of college upperclassmen marked "stubbornness" as wrong but 66 per cent of older men, and 36 per cent of young women but 60 per cent of the older. Similarly for "faultfinding," percentages for young and older men are 59 and 86, and for women 64 and 88; for "temper" they are 25 and 55, 31 and 58; for "fretting" 41 and 60, 28 and 51. In short, the many percentages, at first confusing, upon close study seem in diverse ways intelligible.

The lower part of Table 1 shows college students more often worried about "marriage," also "looks" and "popularity," in '53 than in '23; this finding might be a product of earlier marriage now and increased proportion of college women marrying. That the older group should worry less about these items would be expected—and that older men would worry more about business than students. The last five lines of the table are again items with low total difference because likely to be worried about more or less at any age and any period.

That 28 per cent of young men worried about "health" but 41 per cent of the older was an increase to be expected; the analogous figures for young

<sup>3</sup>The graph shows condemnation from the grades on. Graphs for a number of these items through '43 will be found in reference 6. And any interested reader may obtain from the writer mimeographed sheets giving all 10,000 percentages of the total item analyses, also ozalid copies of a few graphs not included in this paper.



and older women are 31 and 41. The analogous figures for "disease" are 27 and 48, 30 and 38; for "weakness," 37 and 41, 33 and 46; for "dizziness," 9 and 22, 14 and 25; for "falling," 21 and 22, 20 and 22; for "helplessness," 20 and 18, 21 and 36; for "death," 21 and 26, 48 and 21. The above items illustrate how, scattered through the tests, diverse items having interesting bearings on a given topic may be found; and the percentages show that many young people in their prime are anxious about health while the older are not, comparatively, over-concerned.

The oldest group seems also not religiously disturbed, nor the younger generation without conscience. Worries of '23 and '53 junior-seniors and of older adults in '53 for these groups in order, men and women, ran as indicated in percentages for certain items in this general category (Table 2).

TABLE 2

	Men			Women		
religion	27	28	15	38	42	18
God	36	29	14	28	33	5
conscience	36	26	29	39	38	39
morals	38	33	41	33	48	39
sin	55	33	42	43	38	31

Here there seems to be no clear difference between '23 and '53, but the older seem to worry less about the first two topics. And it should be mentioned that these words (like other aggregates brought together for discussion) are scattered all through the "worry" test—"sin" in the second line, "God" in the eighth, "conscience" in the twelfth, and so on; the similarities in the percentages are not due to these items appearing together on the test sheet.

On the interest test, not given to the adults, the one basic comparison of the present study was of '23 and '53 junior-seniors. The test may well be called an odd conglomerate of topics. The writer still recalls the half-humorous curiosity with which he assembled the material—mostly during a vacation—some 35 years ago, with the purpose to get a wide-ranging quick check of the interests of children and young adults. For the items which seem of most interest, Table 3 gives '23 and '53 percentages for men, then women.

Little comment seems necessary; the grouping given the items indicates certain relationships, some already commented upon. Greater liking for Beethoven might be related to the many present-day symphony broadcasts and classical recordings, and less interest in Raphael to lack of comparable means for popularizing the fine arts. Edison died in 1931. The vogue of dance bands is evidenced, the less excitement about the athletic girl now

TABLE 3

	Men		Women			Men		Women	
	'23	'53	'23	'53		'23	'53	'23	'53
Beethoven	42	59	69	79	kissing	65	89	26	87
Edison	74	42	53	25	flirting	47	63	15	41
Napoleon	62	41	31	15	athletic girls	64	46	60	31
Raphael	37	20	64	32	babies	50	74	76	96
bands	61	74	46	89	children	65	77	82	98
banjos	56	32	50	34	housekeeping	16	27	63	79
					bargains	39	66	46	73
loafing	22	50	7	40	onions	34	51	22	58
talking	63	85	68	95	mustard	29	59	22	56
smoking	61	41	10	46	pop	48	73	39	73
business	78	45	34	37	prayer	31	56	36	68
doctors	63	75	63	86	church	67	76	69	75

than when she was more of a novelty, the consistent growth of sex-social and familial interests. The increased interest in babies and children is in interesting congruence with data showing earlier marriage now, more college girls marrying, and more offspring. The increased liking for onions seems at first odd as well as trivial. But perhaps the older reader will remember being told in childhood that really the best people did not eat that breath-tainting vegetable; now, however, snackbars and hamburgers and hot dogs seem to have popularized onions and mustard. Trivial items perhaps—but genuine pieces of Americana. Finally, the increased interest in prayer and church suggests that the youth of today is serious—and is in accord with recent reports regarding church attendance.

### C. DISCUSSION

The mass of the statistical detail, the unusual nature of the test, and its almost cosmic range of topics touched, at first seem to make any coherent interpretation of the results as a whole almost impossible. But close study yields, it is believed, certain generalizations hardly to be reached with a device less wide-reaching or more conventional.

#### 1. *Mid-West Youth, from '23 to '53*

Most clearly evident are the sex-social changes from 1923 to 1953. Not only are these changes substantial, consistent from one decade to the next and at each age, and from one item to another related item, they involve more than increased freedom in sex-social relationships. There is no increase in approval of divorce, more anxiety about marriage, more liking for babies and children and housekeeping. In general, the changes are greater over the 30-year period for women than men. One might hypothecate that the



fair sex, having won its political and economic and educational equality, is using its increased status and freedom more fully to be itself. Far from being primarily license (as some critics of Kinsey would imply), on the whole, the net outcome would seem healthy.

Increasingly the "blue laws" are disappearing as regards amusements. Even in '23, few college students considered dancing or cards wrong. However, in '23, many grade school children did; if their attitude be considered an indication of their "parents' views, then general acceptance of these diversions was then rather recently achieved. The curves show smoking in process of acceptance from '23 on for women, poolrooms less condemned. But it is believed that more is involved here than simply increased acceptance of certain amusements and minor vices. In '23, the tendency to label a thing "wrong" seemed much more wide-spread. So overeating was then more often considered wrong, not merely unwise or a hazard to health or appearance, and dullness and clumsiness were also marked "wrong," not simply unfortunate. All this is interpreted as involving a basic change in cultural point of view from what might be called a Puritanical to a rational or naturalistic moral philosophy—a less generally condemnatory and more accepting and kindly point of view. But again there is not a relaxing of all standards; thoughtlessness is as generally condemned in '53 as earlier, and interests in prayer and church have increased. Unfortunately, the tests do not include as much matter involving desirable traits as would here be wished. But again it seems not unreasonable to conclude that the educated young person of the present is not merely more uninhibited but more socially responsible than before.<sup>4</sup>

The above findings are for college students. As mentioned in discussion of the first graph, it shows eighth grade children marking a few less items wrong in '43 than in '23 but college seniors marking many fewer wrong on the later date. Both times, the drop from college freshman to senior year is rapid.<sup>5</sup> But in '53, the freshman-senior drop is slight; a projection of the curves back would suggest that, if the tests had been given grade school children also then, they would have considered many fewer items wrong than in '43—that at least young adult attitudes of '23 would have shown more in parental attitudes. Presumably times of cultural change like the past 40

---

<sup>4</sup>To obtain more evidence on positive traits, the admirations test of the "interest-attitude" form (6) was given a few college students. The test asks them to mark those traits they admire in a list of 90 characteristics of personality. This form was given to large numbers of students in '33. The data for 1953 were in various ways unsatisfactory. But they are in congruence with the above conclusion in showing college women then substantially more often than in '33 admiring persons who are cooperative, patient, charitable, kind, fair.



years show especially great changes from home moral teachings to young adult codes—a rate of change might thus roughly be measured. If the X-O tests had been given in 1823 or were to be given in 1963, less child-adult differences than in '23 or '43 might be hypothecated.

## 2. *The Necessity of Comparative Data, for Fair Appraisal of Older People*

The indication that the supposed conservatism of older people might often be a continuance of the attitudes of their youth, rather than an actual increasing conservatism with age, is believed a distinctive contribution of this investigation. Indeed, the data suggest that old persons have really become more liberal on certain issues than they had been when young, being conservative only by comparison with the young now (fewer young women than older now mark smoking wrong, but when these older women were young, many more young women so marked). The substantial percentages of old people worrying about health matters seem less hypochondriacal beside the considerable percentages of young people similarly anxious. The comparatively small number of 60-year-olds worrying about death and religion seems less like indifference or apathy and more like quiet confidence and courage when it is noted that more 40-year-olds worry about both these topics. Fairness to each age calls for comparison with others, and older people should be seen in relation to their own as well as present day youth. Personality schedules intended simply for the study of old people (like the Cavan-Burgess-Havighurst inventory) seem insofar less satisfactory than devices applicable over a wide age range.

## 3. *Distinctive Merits of the Instrument and Method of This Study*

The little X-O Form is indeed thus widely usable; in the present series of studies it has been given from the sixth grade to an old folks home, almost always arousing interest and taken easily. Most present day test batteries and personality inventories are substantial booklets, ostentatiously objective and formal. The opinion is ventured that there might well be more emphasis on simplicity of blank and task—especially in work with adults and older people—and variety.<sup>5</sup>

<sup>5</sup>The X-O includes double yield of score: so on the first test the subject first crosses words thought wrong then circles the one word in each line thought worst (perhaps the first use of a forced choice technique) and the number of lines in which a subject chose other than the modal choice was his "idiosyncrasy" score. More general conformity to the current mode in '53 than '23 and in youth than age was hypothecated—but lack of funds prevented such handling. In 1920, "net differential units" were proposed similar to those Strong has used in differentiating interest in different occupations; and a trial unit to measure maturity of personality was developed. But again lack of funds has prevented further work.



The basic design of the investigation here reported has been the giving at 10-year intervals of the same inventories in the same place to comparable groups. The form used was not planned originally for such purpose. A battery which was so planned, including perhaps units concerned with basic abilities, elements of schooling, personality traits, given at regular intervals in a variety of places (perhaps Midwest, East, South, Hawaii, perhaps abroad—the X-O is currently being given in India) and to a variety of age groups, could yield data of cumulatively great importance. It would seem a project to which a foundation-financed continuing committee might well address itself.

#### D. SUMMARY

The paper reports the fourth in a series of decade surveys begun in 1923 with the X-O Form—a simple folder on which individuals check items they think wrong such as smoking, about which they worry such as money, or which they like as traveling. In the present study, some 800 college undergraduates at Ohio State University, plus 400 adults ranging in age from the twenties to the seventies, took the tests.

A decrease from 1923 to 1943 in number of borderline acts (such as smoking) considered wrong was found to continue in 1953, and was especially marked for women, in the college years. Older adults marked more items wrong than did the younger; the older marked about the same number as young people in the decades when these adults were young. These adults thus appeared not so much to show increasing conservatism with age as a retention of the attitudes of their youth. Worries showed little change in total from youth to age and from 1923 to 1953. This finding was interpreted in part as meaning that worries are not part of the changing general culture pattern; in part it might be assumed that the test did not include items now of major concern (as the atom bomb) nor of special concern to older adults. Interests were found to have increased in number since 1923, and especially for the girls. These increases were understood as resulting from increasing leisure time and leisure time interests; and especially a freeing of girls for more diverse interests and recreations.

Item analysis showed decreasing condemnation of and increasing liking for certain freedoms in sex-social relationships, and social amusements; but the trends are so broad as to indicate a general decrease in social taboos and warming social climate. There seemed no lessening in social responsibility. Far from being more conservative than when they were young, the oldest group showed some liberalizing of attitudes.

It is urged that there should be more (a) simple wide-covering inven-

tories and like devices, and (b) regular, systematically planned surveys, as instruments and methods in social psychology.

## REFERENCES

1. ALLEN, F. L. *The Big Change*. New York: Harper, 1952.
2. BUCK, W. A measurement of changes in attitudes and interests of university students over a ten-year period. *J. Abn. & Soc. Psychol.*, 1936, **31**, 12-19.
3. CHAMBERS, O. R. Measurement of personality traits. In *Research Adventures in University Teaching*. Bloomington, Illinois: Public School Publishing Company, 1927. (Pp. 71-80).
4. KINSEY, A. C., *et al.* *Sexual Behavior in the Human Female*. Philadelphia: Saunders, 1953.
5. PRESSEY, S. L. A group scale for investigating the emotions. *J. Abn. Psychol.*, 1921, **16**, 55-64.
6. PRESSEY, S. L., & PRESSEY, L. C. Development of the interest-attitude tests. *J. Appl. Psychol.*, 1933, **17**, 227-38.
7. PRESSEY, S. L. Changes from 1923 to 1943 in the attitudes of public school and university students. *J. of Psychol.*, 1946, **21**, 173-188.

*Department of Psychology*  
*Ohio State University*  
*Columbus 10, Ohio*